Human Immunodeficiency Virus Evasion of APOBEC3 Restriction Factors

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

John Squire Albin

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

Advisor: Reuben S. Harris

October 2012

Prior to my arrival in Minnesota, there were two individuals in particular at the University of Kansas who were especially important to my chosen career path. Dr. Matthew Buechner's teaching inspired my decision to major in microbiology at a time when, having just spent two years as a music major and with next to no scientific background, all I really wanted to do was fulfill the basic pre-med requirements and become a primary care provider somewhere in Latin America. It was also Dr. Buechner who, recognizing my interest in RNA interference, directed me to Dr. Lisa Timmons, whose lab provided the entirety of my undergraduate research experience and kept me thinking about the principles that still underlie how I view the intersection of science and medicine.

My decision to pursue medicine at all in college, the decision that somehow snowballed into this dissertation, was primarily motivated by my interest in HIV. For that, I thank the Diocese of Lincoln and the movie *Kids*, whose cautionary tales planted the fear that became my curiosity and desire to subject myself to thousands of hours of controlled exposure. From this same time period, thanks also go to all of the music teachers who taught me to pipette before I knew what a pipette was.

Before that, I of course owe many of the characteristics that make me who I am to my family. In particular, my parents were largely responsible for cultivating my independence and work ethic. My uncle Tom Albin was especially important for keeping science in the back of my mind via mammoths, simple generators and time travel. The interest in learning necessary to continue on my path comes from many places, but among those that stick out are the endless shelves of books in Bellwood. My openness to biology in particular was in no small way influenced by catching toads and playing with collie litters on my grandparents' farm in the summers. Thanks are also in order for the most recent addition to my family, my wife Shoko, whose love, support and understanding have helped me to retain just enough sanity to power through the end of my PhD. The informed consent process concerning the remainder of my training has yet to faze her.

Dedication

To Coast to Coast AM, whose wild tales of aliens, cryptids, dark conspiracies and assorted paranormal encounters made evolving superviruses in the dead of night fun. The truth is out there.

Abstract

The human immunodeficiency virus accessory protein Vif protects the viral genome from the mutational activity of APOBEC3 subfamily DNA cytosine deaminases by facilitating their proteasomal degradation, thereby preserving viral infectivity. A comprehensive understanding of the components of the Vif-APOBEC3 interaction is therefore important for consideration of the potential for novel antiretroviral approaches aimed at modulating this critical host-pathogen interaction. Here, we establish APOBEC3F among the seven subfamily members as a valid model for the study of the APOBEC3-Vif interaction. By utilizing this model as a starting point, we further define the APOBEC3-Vif interaction sites in each protein and the downstream ubiquitin acceptor sites modified en route to APOBEC3 degradation, in the process deriving broader insights into the nature of the interactions between different APOBEC3 proteins and Vif. In contrast with the diversiform APOBEC3-Vif interactions proposed in the extant literature, we find that the interaction of Vif with different APOBEC3 proteins likely proceeds through a conserved helix-helix interaction. Even if one were to successfully block this interaction for therapeutic purposes, however, the virus may develop accessory mechanisms of APOBEC3 evasion to bypass the intervention. While we find that this can occur, present evidence suggests that such alternatives may be insufficient to circumvent restriction in cells that naturally express multiple APOBEC3 proteins. Thus, it may be possible to potentiate the action of multiple endogenous antiretroviral proteins to counteract human immunodeficiency virus infection by targeting a conserved interaction motif as described herein.

Table of Contents

List of Tables

List of Figures

Author Contributions

The work described in this dissertation is comprised not only of my own efforts, but also those of many other individuals encountered in the course of my doctoral training. This section describes the specific contributions of each individual to each chapter.

Chapter 1: Introduction

John S. Albin – Wrote initial drafts, including all tables and figures, of each of the two manuscripts contained in **Chapter 1**.

Reuben S. Harris – Edited initial drafts, tables and figures.

Chapter 2: Long-term Restriction by APOBEC3F Selects Human Immunodeficiency Virus Type 1 Variants with Restored Vif Function.

John S. Albin – Conceived, designed and executed the project including data associated with all figures save **Figure 2-7**; wrote the manuscript.

Guylaine Haché – Primary author responsible for the precursor to this manuscript who also carried out some of the selections summarized in **Table 2-2**; made some of the stably transfected APOBEC3-expressing cell lines used throughout.

Judd F. Hultquist – Carried out selections associated with **Figure 2-7**.

William L. Brown – Made stably transfected APOBEC3-expressing cell lines for use in these studies.

Reuben S. Harris – Conceived and designed the project while serving as advisor throughout; edited the manuscript.

Chapter 3: A Single Amino Acid in Human APOBEC3F Alters Susceptibility to HIV-1 Vif.

John S. Albin – Conceived, designed and executed the project including experiments associated with **Figures 3-2** through **3-4**; wrote the manuscript.

Rebecca S. LaRue – Designed and executed experiments associated with **Figure 3-1**, prepared **Figure 3-1** and wrote the associated **Materials and Methods**.

Jessalyn A. Weaver – Designed and executed experiments associated with **Figure 3-1**.

William L. Brown – Made the APOBEC3-expressing stable cell lines used in **Figure 3-2**.

Keisuke Shindo – Performed unpublished coimmunoprecipitation experiments confirming the continued physical association of Vif-resistant APOBEC3F with Vif.

Elena Harjes – Made the structural model of the APOBEC3F C-terminus used in **Figure 3-5**, wrote the associated **Materials and Methods** and supplied **Figure 3-6**.

Hiroshi Matsuo – Collaborator in our structural studies of the C-termini of APOBEC3G and APOBEC3F; advisor to Elena Harjes.

Reuben S. Harris – Conceived and designed the project while serving as advisor throughout; edited the manuscript.

Chapter 4: Dispersed Sites of HIV Vif-dependent Polyubiquitination in the DNA Deaminase APOBEC3F.

John S. Albin – Conceived, designed and executed the project; wrote the manuscript.

John S. Anderson – Assisted with experimental design and execution throughout as well as composition of the manuscript.

Jeffrey R. Johnson – Carried out mass spectrometry experiments associated with **Figure 4-3** and wrote the associated **Materials and Methods**.

Elena Harjes – Made the structural model of the APOBEC3F N-terminus used in **Figure 4-4** and wrote the associated **Materials and Methods**.

Nevan J. Krogan – Collaborator in the application of mass spectrometry to questions associated with the interactions between Vif and APOBEC3 proteins; advisor to Jeffrey R. Johnson.

Hiroshi Matsuo - Collaborator in our structural studies of the C-termini of APOBEC3G and APOBEC3F; advisor to Elena Harjes.

Reuben S. Harris – Conceived and designed the project while serving as advisor throughout; edited the manuscript.

Chapter 5: Evidence for a Conserved Structural Feature in HIV-1 Vif That Interacts with APOBEC3 Proteins.

John S. Albin – Conceived, designed and executed the project; wrote the manuscript.

Eric W. Refsland – Responsible for data not shown extending the observations reported here to A3H.

Terumasa Ikeda – Assisted with processing and sequencing of isolates associated with **Figures 5-12** and **5-13**.

John R. Holten – Assisted with processing and sequencing of isolates associated with **Figures 5-12** and **5-13**.

John S. Anderson – Contributed to work not shown extending the observations reported here to A3C and A3D.

William L. Brown – Made the stably transfected APOBEC3-expressing cell lines used throughout.

Reuben S. Harris – Conceived and designed the project while serving as advisor throughout; edited the manuscript.

Chapter 6: Vif-independent Adaptation of HIV-1 to Human APOBEC3 Proteins.

John S. Albin – Conceived, designed and executed the project; wrote the manuscript.

Reuben S. Harris – Conceived and designed the project while serving as advisor throughout; edited the manuscript.

Chapter 7: Conclusions and Discussion

John S. Albin – Wrote the chapter for use in this dissertation.

Reuben S. Harris – Edited the chapter.

Supplementary Chapter 1: Additional Mutagenesis of the APOBEC3F C-terminal Deaminase Domain.

John S. Albin – Conceived, designed and executed the experiments; wrote the supplementary chapter.

Reuben S. Harris – Conceived and designed the experiments while serving as advisor throughout; edited the supplementary chapter.

Supplementary Chapter 2: Catalytic Activity of APOBEC3F Is Required for HIV-1 Restriction.

John S. Albin – Conceived, designed and executed the experiments; wrote the supplementary chapter.

William L. Brown – Made stable APOBEC3-expressing cell lines used throughout.

Reuben S. Harris – Conceived and designed the experiments while serving as advisor throughout; edited the supplementary chapter.

Preface

Human immunodeficiency virus type 1 (HIV), the causative agent of acquired immunodeficiency syndrome (AIDS), has sparked a global pandemic of staggering proportions. In 2011, 2.5 million new infections against 1.7 million deaths brought the total number of people living with HIV to more than 34 million on top of approximately 30 million dead since 1981 (UNAIDS, 2010a, b, 2012). While advances in antiretroviral therapy have commuted the automatic death sentence imposed by the virus in the early years of the pandemic, HIV infection remains fundamentally incurable. Approaches to managing the associated disease burden now focus on three main areas - preventing infection, extending extant therapies to those who cannot readily access them and searching for new methods to combat the virus.

As in all of medicine, prevention is the cornerstone of good health. When prevention fails, current antiretroviral regimens can effectively quarantine the virus indefinitely, but these drugs must be taken continuously and are limited both by their side-effects and by the ability of HIV to acquire resistance mutations that may render a given drug class ineffective. Given these limitations, extensive resources are still directed toward basic research aimed at the discovery of new therapeutic options that might increase and improve the strategies available to clinicians or, in the ideal scenario, proffer a path to a cure.

Among these discovery-oriented endeavors, one of the most promising breakthroughs of the last decade has been the description of host restriction factors including APOBEC3G, TRIM5α, BST-2/tetherin and SAMHD1 [reviewed in (Harris et

1

al., 2012; Malim and Bieniasz, 2012)]. These host restriction factors are cellular proteins that have innate antiviral activities that must be counteracted by HIV to facilitate viral spread, as opposed to host dependency factors, which are cellular proteins that the virus requires to complete its lifecycle. APOBEC3G is a DNA cytosine deaminase that introduces massive levels of mutation into the HIV genome that render the virus nonfunctional in the absence of the virus accessory protein Vif, which degrades APOBEC3G to prevent its antiviral functions (see **Chapter 1**). TRIM5α is a tripartite motif-containing family protein that recognizes the retroviral capsid, interrupting viral replication by stimulating premature uncoating and enhanced immune recognition, the effects of which can be counteracted by simple mutations in the viral capsid that ablate binding of TRIM5α (Pertel et al., 2011; Sayah et al., 2004; Stremlau et al., 2004; Towers et al., 2003). Mechanistically removed from the reverse transcription-centric nature of the first two factors, BST-2/tetherin is a GPI-anchored transmembrane protein that literally tethers HIV and other enveloped viruses to the cell surface to prevent their release and is counteracted by the degradative functions of the viral accessory protein Vpu - or, in related viruses, the viral Env or Nef proteins [(Neil et al., 2008; Van Damme et al., 2008) and reviewed in (Douglas et al., 2010)]. Finally, the latest restriction factor to be described is SAMHD1, a deoxynucleoside triphosphate triphosphohydrolase that maintains deoxynucleoside triphosphate levels in myeloid lineage cells at levels incompatible with viral replication and is counteracted by the lentiviral Vpx accessory protein (Berger et al., 2011a; Goldstone et al., 2011; Hrecka et al., 2011; Laguette et al., 2011; Powell et al., 2011).

Throughout evolutionary history, then, there has existed a conflict between retroviruses and the host defense proteins aimed at controlling their replication, many of which likely remain undiscovered. It is not known how these defenses evolved or how important they have been to the control of exogenous viruses over time, but they may be maintained, at least in part, for the purpose of controlling the endogenous retroelements most commonly encountered by a cell - those comprising half of its own genome. As stated above, however, these proteins collectively carry therapeutic promise since a retrovirus must carry out some action in order to counteract them, generally via the functions of accessory proteins. Where there is an interaction, then, there is a potential novel antiretroviral therapeutic target.

This dissertation focuses on the potential for interrupting the degradation of APOBEC3 subfamily proteins by the viral accessory protein Vif, which in turn may liberate APOBEC3 proteins to carry out their antiretroviral functions. In considering this possibility, it is important to make two key distinctions between APOBEC3 proteins and all extant drugs and known restriction factors. First, there are at least six APOBEC3 proteins in addition to APOBEC3G, all of which have been shown under one condition or another to have anti-HIV activity (see **Chapter 1**). Effectively neutralizing Vif, then, would theoretically result in the potentiation of multiple antiretroviral proteins, yielding a sort of endogenous combinatorial antiretroviral therapy. Second, the mechanism of APOBEC3 restriction involves the introduction into the HIV genome of mutation levels incompatible with survival and spread. Once this genetic information is scrambled, it is difficult if not impossible for the virus to recover what was lost, and there are many such examples of viruses that have become "endogenized" in genomes over time through loss

of function [reviewed in (Stoye, 2012)]. APOBEC3 proteins do not just inactivate or neutralize a viral particle like all other known antiretroviral drugs and naturally occurring restriction factors - they destroy the genome that makes the virus what it is. As such, it is at least hypothetically possible that bombarding HIV with these innate defense proteins may represent not only a novel therapy, but also a potentially curative approach. In considering this dissertation, it is my hope that the reader will recognize the basis for believing that such things, while hardly imminent, are possible.

Chapter 1: Introduction

FOREWORD

The text comprising this chapter is an updated amalgamation of the following references reproduced in sections with permission:

Albin JS and Harris RS. Interactions of host APOBEC3 restriction factors with HIV-1 *in vivo*: implications for therapeutics." *Expert Reviews in Molecular Medicine*. 2010; 12 (e4): 1-26

Albin JS and Harris RS. APOBEC3 proteins and their roles in HIV proviral DNA synthesis. Chapter 14 in Human Immunodeficiency Virus Reverse Transcriptase: a 25-year Success Story. Ed. SF LeGrice, *in press*

Overview of APOBEC3 Proteins and Vif

The human genome has a total of seven *APOBEC3* DNA cytosine deaminase genes arrayed on chromosome 22 (*A3A*, *A3B*, *A3C*, *A3D*, *A3F*, *A3G* and *A3H*), the result of a series of tandem duplications that have expanded the locus throughout primate evolution (Conticello, 2008; LaRue et al., 2009; LaRue et al., 2008). This expansion, combined with evidence for strong positive selection among most of these *APOBEC3* genes, suggests that a primary function of their protein products is defense of the genome against any number of retroelements encountered throughout evolution (Duggal et al., 2011; OhAinle et al., 2006; Sawyer et al., 2004). The list of retroelements for which direct evidence of APOBEC3 restriction exists currently includes endogenous LTR and non-LTR retrotransposons as well as exogenous members of the *Hepadnaviridae* and *Retroviridae* families [*e.g.* (Bogerd et al., 2006b; Derse et al., 2007; Esnault et al., 2005; Esnault et al., 2006; Harris et al., 2003a; Löchelt et al., 2005; Mangeat et al., 2003; Muckenfuss et al., 2006; Navarro et al., 2005; Okeoma et al., 2007; Russell et al., 2005; Sheehy et al., 2002; Stenglein and Harris, 2006; Turelli et al., 2004)].

Chief among the examples of viruses restricted by APOBEC3 proteins is the prototypical lentivirus human immunodeficiency virus type 1 (HIV-1, hereafter abbreviated HIV). In addition to *gag*, *pol* and *env* genes, lentiviruses possess several accessory genes that have integral roles in viral pathogenesis [recently reviewed by (Gramberg et al., 2009; Malim and Emerman, 2008)]. In HIV, these include two accessory genes involved in the regulation of viral gene expression and four additional accessory genes more directly involved in pathogenesis. One of these is *vif,* the virion infectivity factor, which is located in the center of the HIV genome overlapping the 3'

end of *pol* and the 5' end of the accessory gene *vpr*. Expressed from a partially spliced, Rev-dependent subgenomic mRNA (Garrett et al., 1991; Schwartz et al., 1991), the protein product is 192 residues, highly basic and 23 kDa in size (Kan et al., 1986; Lee et al., 1986; Sodroski et al., 1986).

HIV Vif and the Discovery of APOBEC3 Proteins

Vif is required for productive infection *in vivo* and for infection of primary CD4[+] T cells, monocytes, and macrophages *ex vivo* (Desrosiers *et al.*, 1998; Gabuzda *et al.*, 1992; Gabuzda *et al.*, 1994; Strebel *et al.*, 1987). This requirement, however, is variable in some human T cell lines (Fisher *et al.*, 1987), allowing these lines to be grouped as 'permissive' for the replication of Vif-deficient HIV (*e.g.* CEM-SS, SupT1) or 'nonpermissive' (*e.g.* CEM, H9). When produced in permissive cells, virions from Vif-deficient HIV have no quantitative or qualitative defects relative to wildtype virus except for the absence of the Vif protein itself. (Fisher *et al.*, 1987; Strebel *et al.*, 1987). When produced in nonpermissive cells, however, Vif-deficient HIV virions rarely complete reverse transcription (Courcoul *et al.*, 1995; Goncalves *et al.*, 1996; Simon and Malim, 1996; Sova and Volsky, 1993).

Given the above, two scenarios were considered likely to explain the phenotype of Vif-deficient viruses on nonpermissive cells: (1) permissive cells contain an endogenous factor that complements the function lacking in a Vif-deficient virus or (2) nonpermissive cells contain a Vif-suppressible factor with antiviral restriction activity. To determine which hypothesis was correct, two groups tested the infectivity of virions produced from heterokaryons formed by the fusion of permissive and nonpermissive cells (Madani and Kabat, 1998; Simon *et al.*, 1998). Under these conditions, the production of infectious virions would support the former hypothesis, while the production of noninfectious virions would support the latter. In fact, such heterokaryons formed from the fusion of permissive and nonpermissive cells produce virions with diminished

8

infectivity, indicating that nonpermissive cells contain a Vif-sensitive dominant restriction factor.

To identify the cellular factor suggested by cell fusion experiments, Sheehy *et al.* sought to isolate genes uniquely expressed in a nonpermissive cell line called CEM (Sheehy *et al.*, 2002) by suppression subtractive hybridization, which simultaneously suppresses the recovery of transcripts common to CEM and its permissive derivative CEM-SS and enriches the recovery of transcripts that are more abundant in CEM. Among many sequences, a cDNA encoding APOBEC3G (A3G) was isolated, and engineering naturally permissive CEM-SS cells to express A3G rendered these cells nonpermissive for the replication of Vif-deficient but not wildtype virus. A3G therefore accounts for the nonpermissive phenotype of CEM, although a number of studies since then suggest that other APOBEC3 proteins may also contribute to the nonpermissive phenotype (**Table 1-1**). Of particular note, passage of Vif-deficient HIV on permissive T cells stably transfected with A3G yields viral variants that grow efficiently in the presence of A3G but not in CEM cells (Haché *et al.*, 2008), implying that other factors beyond A3G are also capable of restricting HIV in nonpermissive cells.

At approximately the same time, two other groups independently identified A3G, not for its antiviral activity but rather for its enzymatic activity, which provided a crucial clue about the mechanism by which A3G might restrict HIV. Jarmuz *et al.* identified A3G as part of a family of polynucleotide cytosine deaminases related to the mRNA-editing enzyme APOBEC1, which edits the apolipoprotein B mRNA (Jarmuz *et al.*, 2002). The acronym APOBEC derives from APOBEC1 and stands for 'apolipoprotein B mRNA editing, catalytic polypeptide'; human APOBEC3 proteins are designated

'apolipoprotein B mRNA editing, catalytic polypeptide-like 3' followed by the letter, A–H, identifying the specific protein. Meanwhile, Harris and colleagues hypothesized, based on homology between APOBEC1 and the DNA mutator activation-induced deaminase (AID), which is involved in antibody diversification, that APOBEC1 and its homologues can act on DNA substrates. Indeed, APOBEC1, A3G, A3C and AID were all shown to mutate DNA (Harris *et al.*, 2002; Petersen-Mahrt *et al.*, 2002).

Putting the enzymatic and restrictive abilities of A3G together, several groups soon thereafter established that A3G restricts HIV by introducing massive numbers of C-to-U transitions in the viral minus strand cDNA during reverse transcription in target cells (Harris *et al.*, 2003a; Lecossier *et al.*, 2003; Mangeat *et al.*, 2003; Zhang *et al.*, 2003). These transitions may then either trigger the degradation of viral cDNA or impede reverse transcription (see below) to decrease overall reverse transcript accumulation or they may become fixed, manifesting as plus-strand G-to-A hypermutation that yields a presumably hypofunctional mutant virus (**Figure 1-1A**). To prevent this from happening, Vif links A3G in producer cells to an E3 ubiquitin ligase complex, thus facilitating its proteasomal degradation and permitting the production of relatively A3G-free virions that can successfully complete reverse transcription in target cells (**Figure 1-1B**).

While discussion to this point has focused on a basic description of the interactions between Vif and A3G, all of the six additional proteins in the APOBEC3 subfamily – A3A, A3B, A3C, A3D, A3F and A3H – have been implicated at some point in the mutation and/or restriction of HIV under certain conditions. A summary of references with data concerning the restriction activity and Vif sensitivity of each APOBEC3 protein is provided in **Table 1-1**.

While the results apparent in **Table 1-1** do reflect a lack of consensus on the ability of certain APOBEC3 proteins to function against HIV, there are areas of agreement. Most notably, it is well accepted that A3F and, more prominently, A3G have strong activity against HIV, although a pair of recent reports have impugned the ability of stably expressed A3F to restrict the virus (in contrast with **Chapters 2**, **3**, **5** and **6** of this thesis as well as **Supplementary Chapter 2**), leaving A3G as the only universally agreed upon APOBEC3 antagonist of HIV (Miyagi *et al.*, 2010; Mulder *et al.*, 2010). Beyond A3F and A3G, there is a second tier consisting of A3B and A3D where antiviral activity is generally reproducible in overexpression single-cycle infectivity assays but may not be apparent in a more physiologic spreading infection system. Recent results from our lab have supported a minor role for A3D in the restriction of HIV and a somewhat more robust role for this APOBEC3 protein in the hypermutation of the virus, but these studies failed to note an effect of A3B on HIV outside of 293T overexpression systems (Hultquist *et al.*, 2011; Refsland *et al.*, 2012). That said, the fact that A3B expression increases significantly in lymphoid tissue from acutely infected patients along with A3F and A3G (Li *et al.*, 2009) combined with the previously reported correlation between A3B haplotypes and clinical indicators of disease may also suggest a role for this Vif-insensitive protein *in vivo* [(An *et al.*, 2009) and see below].

A3H represents a special case since only two known allelic variants appear to encode a stable protein, and stability is required for its ability to effectively restrict HIV (Harari *et al.*, 2009; OhAinle *et al.*, 2008; Tan *et al.*, 2009; Wang *et al.*, 2011a). A3C has historically been something of an enigma among the APOBEC3 proteins, well-expressed in many cells and tissues but with no consistent phenotype of any kind, let alone a

11

reliable restrictive effect against HIV, despite its exquisite Vif sensitivity [(Refsland *et al.*, 2010) and **Table 1-1**]. It should be noted, however, that A3C is quite active against simian immunodeficiency virus (Yu *et al.*, 2004a).

In an attempt to bring some order to the often contradictory reports listed in **Table 1-1**, our lab has recently published two studies assessing the restriction capacity of all seven APOBEC3 proteins in T cells, which offers the advantage of direct comparison in consistent systems as opposed to a patchwork of many systems in many hands with many divergent goals. By either stable expression of each of the APOBEC3 proteins in APOBEC3-deficient T cell lines or by the knockout or knockdown of endogenous APOBEC3 proteins in naturally nonpermissive cells, these studies have converged on four APOBEC3 proteins as the components of the repertoire relevant to the restriction and hypermutation of HIV in T cells – A3D, A3F, A3G and A3H (Hultquist *et al.*, 2011; Refsland *et al.*, 2012).

While some case can thus be made for the ability of six of seven human APOBEC3 proteins to dominantly restrict HIV in T cells, particularly A3D, A3F, A3G and A3H, there are few findings to support such a role for A3A. The caveat to this is that some authors have proposed A3A as a source of sequence diversification in HIV given its expression and apparent mutational activity in myeloid lineage cells (Koning *et al.*, 2011; Koning *et al.*, 2009; Refsland *et al.*, 2010; Thielen *et al.*, 2010). Some have even proposed A3A as the long-sought Vpx-suppressible myeloid cell restriction factor (Berger *et al.*, 2010; Berger *et al.*, 2011b; Peng *et al.*, 2007), consistent with its role in the clearance of foreign DNA from myeloid cells (Stenglein *et al.*, 2010). Nevertheless, there is presently a movement toward acceptance of the deoxyribonucleoside triphosphate

triphosphohydrolase SAMHD1 as "the" Vpx-suppressible suppressor of HIV reverse transcription in myeloid cells (Berger *et al.*, 2011a; Goldstone *et al.*, 2011; Hrecka *et al.*, 2011; Laguette *et al.*, 2011; Powell *et al.*, 2011). Whether A3A and SAMHD1 act in concert remains unknown.

Mechanisms of Restriction

Restriction by APOBEC3 proteins is dependent on their expression in virus producing cells and on their encapsidation during virus assembly [(Harris *et al.*, 2003a; Mangeat *et al.*, 2003; Zhang *et al.*, 2003) and **Figure 1-1A-B**]. This is thought to occur via an association of APOBEC3 proteins with the viral nucleocapsid protein via an RNA bridge (Alce and Popik, 2004; Bogerd and Cullen, 2008; Cen *et al.*, 2004; Luo *et al.*, 2004; Schafer *et al.*, 2004; Svarovskaia *et al.*, 2004; Zennou *et al.*, 2004). In particular, this interaction may be facilitated by the cellular RNA 7SL, which is selectively packaged into HIV virions (Tian *et al.*, 2007; Wang *et al.*, 2007; Wang *et al.*, 2008). The importance of RNA in facilitating the encapsidation of APOBEC3 proteins is further supported by the observation that mutations in the N-terminus of A3G that ablate RNA binding also ablate Gag interaction, encapsidation and, consequently, restriction (Friew *et al.*, 2009; Huthoff *et al.*, 2009; Huthoff and Malim, 2007; Navarro *et al.*, 2005). In total, it is thought that 7 ± 4 molecules of A3G are packaged per particle and that as few as one may be measurably antiviral (Browne *et al.*, 2009; Xu *et al.*, 2007).

Once associated with the viral core, APOBEC3 proteins are positioned to exert their restrictive activities upon nascent reverse transcribed cDNA in target cells. Interpretations of the relative importance of different putative mechanisms of A3G inhibition of HIV may vary, but a broad point of unity in the field is that deamination is a major component of the restriction mechanism, as initially proposed [(Harris *et al.*, 2003a; Mangeat *et al.*, 2003; Zhang *et al.*, 2003) and see *Deaminase-independent Restriction Mechanisms* below as well as **Supplementary Chapter 2**]. Thus, we begin a consideration of the impact of APOBEC3 proteins on proviral DNA synthesis with a

more detailed rendering of the deamination process, focusing on the prototypical APOBEC3 protein A3G, but with the understanding that other APOBEC3 proteins are thought to work similarly.

A3G Deamination and Hypermutation – The polynucleotide intermediates that occur in the course of HIV reverse transcription include the single-stranded RNA genome (ssRNA), a DNA:RNA heteroduplex following synthesis of the minus-strand cDNA templated by the plus-strand RNA genome, single-stranded DNA (ssDNA) resulting from RNase H degradation of the plus-strand ssRNA template after minus-strand synthesis, and double-stranded DNA (dsDNA) after completion of the plus-strand cDNA (**Figure 1-2A-D**). Among these, A3G is capable of binding to both ssRNA and ssDNA (Iwatani *et al.*, 2006; Yu *et al.*, 2004b). Binding to ssRNA is thought to segregate A3G into enzymatically-inactive high molecular mass complexes (Chelico *et al.*, 2006; Wedekind *et al.*, 2006), and this may also be true of association between A3G and viral genomic ssRNA (Soros *et al.*, 2007). In its low molecular mass form, however, A3G is capable of deaminating ssDNA substrates (Harris *et al.*, 2003a; Iwatani *et al.*, 2006; Suspène *et al.*, 2004; Yu *et al.*, 2004b). This ssDNA specificity is reinforced by the observation that there are two gradients of 5' → 3' increasing levels of hypermutation in the viral genome, one from the primer binding site to the central polypurine tract and another from the central polypurine tract to the 3' polypurine tract (Suspène *et al.*, 2006; Yu *et al.*, 2004b). Thus, there is a correlation between the most heavily deaminated regions of the viral genome and those regions of the minus-strand that remain single-stranded for the longest duration during reverse transcription.

Some authors have noted an increased C-to-T transition rate in the viral plus-strand as well, particularly with A3F (Albin *et al.*, 2010a; Bishop *et al.*, 2004; Liddament *et al.*, 2004; Yu *et al.*, 2004b). It is not clear whether these represent a modest ability of A3F to deaminate RNA, RNA:DNA heteroduplex or dsDNA templates or a more mundane phenomenon such as transient ssDNA states within the predominantly double-stranded context in which the plus-strand generally exists. Regardless, there is no evidence for A3G activity on any non-ssDNA substrate, including free nucleosides (Iwatani *et al.*, 2006; Suspène *et al.*, 2004).

While A3G is termed a "deaminase", it actually consists of two conserved deaminase motifs, N- and C-terminal. Between these, only the C-terminus is capable of deaminating ssDNA cytosine to uracil (Haché *et al.*, 2005; Iwatani *et al.*, 2006; Langlois *et al.*, 2005; Navarro *et al.*, 2005; Newman *et al.*, 2005). The catalytically inactive N-terminus, in contrast, is more important for binding to single-stranded nucleic acid substrates, an activity that mediates RNA-dependent oligomerization and encapsidation and that may also help orient A3G for processive deamination (Chelico *et al.*, 2010; Friew *et al.*, 2009; Gooch and Cullen, 2008; Huthoff *et al.*, 2009; Huthoff and Malim, 2007; Li *et al.*; Navarro *et al.*, 2005). Although it has been proposed that monomeric A3G may also be capable of encapsidation, deamination and restriction (Opi *et al.*, 2006), recent evidence suggests the particular mutant employed in this study, while monomeric, may not actually encapsidate, raising the question of whether its apparent ability to restrict represents a specific effect on HIV or a nonspecific poisoning of producer cells [(Friew *et al.*, 2009; Shlyakhtenko *et al.*, 2011) and Mueller & Harris Labs, unpublished data].

Deamination by A3G appears to occur in the 3' → 5' direction (Chelico *et al.*, 2006; Furukawa *et al.*, 2009) and can proceed from 3 to at least 100 nucleotides along ssDNA substrates in the absence of any associated cofactors (Chelico *et al.*, 2006). Although, as stated above, the C-terminal domain is required for actual catalysis, the inactive N-terminal deaminase domain is important for regulating the directionality and processivity of deamination (Chelico *et al.*, 2010). Movement of the enzyme from one region of ssDNA to another may be achieved through some combination of random "jumping" motions to different segments (Chelico *et al.*, 2006) and/or via intersegmental "transfers" in which the simultaneous binding of A3G to distinct ssDNA segments may facilitate its movement from one region to another (Nowarski *et al.*, 2008).

The oligomeric form of A3G that mediates deamination in the context of infection is not entirely clear. Some *in vitro* data have implicated A3G monomers as the enzymatically active form (Nowarski *et al.*, 2008), others dimers (Chelico *et al.*, 2006; Shlyakhtenko *et al.*, 2011). Atomic force microscopy studies that visualize A3G bound to ssDNA substrates have found a range of oligomeric forms, with a predominance of monomers and higher order oligomers (Chelico *et al.*, 2010; Chelico *et al.*, 2008), although more recent evidence derived from a refinement of this technique suggests a model in which unbound A3G, in agreement with the aforementioned prior studies, is predominantly monomeric, while bound A3G can take on a dimeric form with substantial components of monomers and higher order oligomers also apparent (Shlyakhtenko *et al.*, 2011). Although oligomerization of A3G does not appear to be required for deaminase activity (Chelico *et al.*, 2010), the existence of these oligomeric forms may promote

enzyme processivity (Chelico *et al.*, 2008) or the ability of A3G to "jump" or "transfer" to new substrates (Feng and Chelico, 2011; Nowarski *et al.*, 2008).

Different APOBEC3 proteins exhibit different dinucleotide preferences for deamination, with A3G preferring 5'-C<u>C</u>-3' dinucleotides (5'-<u>G</u>G-3' on the plus-strand), often within a broader 5'-CC<u>C</u>A-3' context (5'-T<u>G</u>GG on the plus-strand) [*e.g.* (Beale *et al.*, 2004; Harris *et al.*, 2003a; Harris *et al.*, 2002; Mangeat *et al.*, 2003; Yu *et al.*, 2004b; Zhang *et al.*, 2003)], and other APOBEC3 proteins including A3F typically showing varying degrees of preference for 5'-T<u>C</u>-3' motifs (5'-<u>G</u>A-3' on the plus strand) [*e.g.* (Bishop *et al.*, 2004; Hultquist *et al.*, 2011; Liddament *et al.*, 2004; Zheng *et al.*, 2004) and reviewed in (Albin, 2010)]. This complicates the attribution of hypermutation patterns to different APOBEC3 proteins. That is, while 5'-<u>G</u>G-3' context mutations are most likely attributable to A3G since this is the sole APOBEC3 protein that displays a strong preference for this dinucleotide context, the 5'-<u>G</u>A-3' context mutations could be introduced by any of several APOBEC3 proteins present and potentially active in primary T cells as well as by the myeloid-specific editing activity of A3A (Koning *et al.*; Koning *et al.*, 2009; Refsland *et al.*, 2010; Thielen *et al.*, 2010). Since diverse authors have described substantial amounts of both 5'-<u>G</u>G-3' and 5'-<u>G</u>A-3' context hypermutation in patient-derived viral sequences [reviewed in (Albin, 2010)], the question of which APOBEC3 proteins contribute to the sequence diversification of HIV remains an outstanding and subtly complex issue in the field. Emergent data derived from gene targeting in the naturally nonpermissive cell line CEM2n suggests, however, that A3F and A3D likely contribute to 5'-<u>G</u>A-3' context hypermutation on top of the 5'-<u>G</u>G-3' context hypermutation characteristic of A3G activity (Refsland *et al.*, 2012).

The structural components of A3F and A3G that guide these context preferences have been mapped to loop sequences between the β4 strand and the α4 helix of active deaminase domains, which is thought to interact with ssDNA sequences 5' of the target cytosine (Carpenter *et al.*, 2010; Conticello, 2008; Kohli *et al.*, 2009; Kohli *et al.*, 2010; Langlois *et al.*, 2005). The effect of these loop sequences may further be enhanced by a more N-terminal region corresponding to arginine residues along one of the predicted paths of ssDNA bound to the A3G C-terminus, although the exact placement of ssDNA on A3G remains controversial and, fundamentally, unknown (Carpenter *et al.*, 2010; Chen *et al.*, 2008; Holden *et al.*, 2008). An additional study focusing on A3F and A3C further narrowed these determinants primarily to a single amino acid change within this same region (Langlois *et al.*, 2005). Interestingly, these putative sequence recognition motifs are found adjacent to or extending into the conserved α4 helix where most known determinants of APOBEC3 susceptibility to Vif occur, with the most prominent single amino acid determinant of target sequence specificity in A3F also corresponding to the equivalent of A3G N-terminal amino acid D128, the best-known of these determinants of Vif susceptibility (Albin *et al.*, 2010b; Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Smith and Pathak, 2010; Xu *et al.*, 2004; Zhen *et al.*, 2010). Thus, while the ability of different deaminase domains among diverse APOBEC3 proteins to mediate distinct processes – Vif binding, RNA binding, sequence recognition, etc. – implies that the structure of each will be somewhat different, determining the structure of the C-terminus of A3F may be particularly informative for understanding how this critical region influences the antiviral properties of APOBEC3 proteins and may

explain why some authors have noted a direct effect of Vif on the deaminase activity of A3G (Britan-Rosich *et al.*, 2011; Santa-Marta *et al.*, 2005).

On the other side of this interaction, a study utilizing nucleoside analogues as molecular probes for the DNA determinants recognized by A3G during the deamination process has suggested several important features of the target deamination hotspot itself beyond the aforementioned sequence specificity observations (Rausch *et al.*, 2009). Insertion of an abasic site or alteration of the 2' deoxyribose within a deamination hotspot are both potent inhibitors of deamination at a target cytosine, and ring positions 3 or 4 of the nucleotide base 5' of the target cytosine may be particularly important for hydrogen bonding with A3G residues involved in substrate recognition. These data are consistent with the exquisite specificity of A3G for DNA substrates as well as the importance of active site aromatic residues for enzyme function (see below).

To date, seven structures of the catalytically active C-terminus of A3G have been solved by either NMR or X-ray crystallography (Chen *et al.*, 2008; Furukawa *et al.*, 2009; Harjes *et al.*, 2009; Holden *et al.*, 2008; Li *et al.*, 2012; Shandilya *et al.*, 2010). These structures show that A3G has a concave active site with a floor containing critical hydrophobic residues likely involved in interaction with substrate bases, as also suggested by the above nucleoside analogue studies. Surrounding the outer edges above this concave active site are a number of positively charged residues that may orient ssDNA within the active site, although the exact path taken by ssDNA remains a point of controversy with competing "brim" (Chen *et al.*, 2008) and "kink" (Holden *et al.*, 2008) interaction models proposing somewhat different DNA binding paths. Barring a co-crystal structure, ideally including the full-length enzyme, this debate may be difficult to

resolve, since molecular dynamics simulations have suggested that the DNA binding loops are flexible and that substrate binding may induce a conformational change in A3G (Autore *et al.*, 2010).

cDNA Degradation – In addition to the mutagenic implications of A3G action, early models of restriction proposed that A3G might trigger the degradation of viral cDNA through a base excision repair pathway [(Harris *et al.*, 2003a; Harris *et al.*, 2003b; Mangeat *et al.*, 2003; Zhang *et al.*, 2003) and **Figure 1-2A**]. Potential degradation was considered an important component of any explanatory model because prior observations had indicated that decreases in cDNA accumulation to roughly half of wildtype among early reverse transcription products with still more potent effects on late reverse transcription products was the primary phenotypic difference between Vif-deficient and - proficient viruses (Sova and Volsky, 1993; von Schwedler *et al.*, 1993). In this model, the presence of DNA uracils introduced by A3G would cause a uracil DNA glycosylase enzyme to excise the offending uracil, yielding an abasic site that would then be a target for cleavage by an AP endonuclease (Harris *et al.*, 2003a; Harris *et al.*, 2003b; Mangeat *et al.*, 2003; Zhang *et al.*, 2003). The first critical step in this proposed degradative pathway, uracil excision, is easily envisaged given the fact that UNG2 is packaged into HIV virions, and there is some evidence that this is important for the ability of HIV to repair uracil-containing cDNA (Chen *et al.*, 2004; Mansky *et al.*, 2000; Priet *et al.*, 2005). Indeed, sequencing of proviruses with identical integration sites but different subsets of G-to-A mutations suggests that some uracilated minus-strand cDNAs do undergo repair, although the timing of this is unknown (Yu *et al.*, 2004b).

21

In contrast, others have found that the presence of UNG2 in virions may be detrimental to virus infectivity, presumably via cDNA degradation at the same abasic sites thought to lead to DNA repair in the aforementioned studies (Schröfelbauer *et al.*, 2005; Yang *et al.*, 2007a). This view is not universally held, as others have reported that the presence of UNG2 or another uracil DNA glycosylase, SMUG1, is simply irrelevant to viral infectivity (Kaiser and Emerman, 2006; Langlois and Neuberger, 2008; Mbisa *et al.*, 2007; Schumacher *et al.*, 2008). It has even been proposed, somewhat counterintuitively, that heavy uracilation of viral cDNA may actually represent an adaptive strategy by which HIV can prevent autointegration (Yan *et al.*, 2011).

Setting aside the role of UNG2 in determining the fate of uracilated cDNA for a moment, a separate but related question is what endonuclease(s) might cause the ultimate degradation of cDNA containing abasic sites as proposed. APE1 has been found to be packaged in virions along with UNG2 in a scenario perfectly supporting the base excision repair model of cDNA degradation as originally proposed, but we are unaware of any additional reports confirming this effect (Yang *et al.*, 2007a). It is not clear that an active degradation mechanism would be necessary to result in decreased cDNA accumulation, however, as the mere existence of a number of abasic sites in the absence of a second strand to facilitate repair may simply impede the ability of reverse transcriptase to synthesize a second strand (**Figure 1-2C**). Reverse transcriptase is capable of inserting dAMP across from an abasic site, a phenomenon that may further be stimulated by Vif (Cancio *et al.*, 2004), but this process appears to also reduce enzyme processivity (Cai *et al.*, 1993). Thus, while it is clear that a fundamental aspect of A3-dependent restriction is the decreased accumulation of cDNA, whether this is the result of cDNA degradation,

deaminase-dependent decreases in reverse transcript synthesis, or an alternative mechanism such as one of the deaminase-independent models to be discussed below remains an open and interesting question.

Summary of Deaminase-Dependent Mechanisms – The deaminase-dependent mutation of the HIV genome is a widely accepted mechanism by which A3G can inhibit the virus. The picture surrounding cDNA degradation, however, is somewhat less clear. It seems the crux of the argument is whether an abasic site is generated during reverse transcription since reverse transcriptase should not distinguish cDNA uracils from thymines (Klarmann *et al.*, 2003).

The ssDNA specificity of APOBEC3 proteins has been taken to imply that the substrate for uracil DNA glycosylase activity should also be ssDNA. This ssDNA limitation, in turn, would narrow the available uracil DNA glycosylase enzymes to those with specificity for single-stranded DNA, namely UNG1/2 and SMUG1 in human cells, and as discussed above, there is little consensus on the involvement of these two enzymes in retroviral restriction. It is important to remember, however, that ssDNA, like ssRNA, is capable of taking on secondary and tertiary structures, sometimes with thoroughly impressive effects [*e.g.* (Breaker, 2004)]. Such cDNA folded back on itself, based on the extensive structure of the ssRNA regions of the viral genome first reverse transcribed [*e.g.* (Watts *et al.*, 2009)], would likely represent an appropriate template for alternative uracil DNA glycosylases with double-stranded specificity such as TDG or MBD4 (Krokan *et al.*, 2002) without invoking the specter of unidentified enzymes, a particularly intriguing possibility given the interaction of MBD4 with relatives of the APOBEC3

enzymes (Rai *et al.*, 2008). The assumption of single-strandedness as depicted in **Figure 1-2A** and in many similar figures in prior reviews may therefore need to be revisited.

If the abasic site forms, it is a short leap to cleavage of the phosphodiester backbone, if such cleavage would in fact be necessary versus simple stalling of reverse transcriptase (**Figure 1-2C**). This could occur via enzymatic means as initially proposed, via APE1 or an alternative enzyme. Spontaneous cleavage may also be a viable but, to date, untested alternative hypothesis. For example, the deamination reaction itself releases ammonia, which may hypothetically create a locally basic environment in the viral core favoring destabilization of an abasic site. One could also imagine the apparent cDNA degradation observed experimentally as an artifact caused by sensitivity of an abasic site to the heat denaturation typical of the PCR procedures that have been utilized universally, to our knowledge, to amplify viral reverse transcripts for further analysis [*e.g.* (Borman *et al.*, 1995; Courcoul *et al.*, 1995; Goncalves *et al.*, 1996; Simon and Malim, 1996; Sova and Volsky, 1993; von Schwedler *et al.*, 1993)]. While we acknowledge that alternative routes to cDNA degradation are highly speculative, we propose that, given the lack of consensus on the exact mechanism by which cDNA accumulation is inhibited, it may be appropriate to consider less obvious answers to this important mechanistic question.

Deaminase-independent restriction mechanisms – While the first papers describing the mechanism of A3G action focused on the implications of deamination for HIV restriction, multiple labs soon thereafter began to note that deaminase-independent mechanisms of restriction may also be at play. This was based on the simple observation

that overexpression of A3G containing mutations of conserved residues critical for deaminase activity such as the zinc coordinating residues and the catalytic glutamate of the conserved HxE/PCxxC motif might not fully ablate restriction (Bishop *et al.*, 2006; Newman *et al.*, 2005; Shindo *et al.*, 2003). Although at odds with identical experiments published in the initial descriptions of A3G's deaminase-dependent restriction activities (Mangeat *et al.*, 2003; Zhang *et al.*, 2003), this has grown to become a point of contention within the field. The details surrounding these proposed deaminase-independent mechanisms fall into several categories, but these studies are united in the suggestion that the decreased accumulation of reverse transcripts characteristic of the nonpermissive phenotype is not due to cDNA degradation, but rather to the inhibition of cDNA synthesis in a manner not dependent on the introduction of uracils or abasic sites into viral cDNA.

Physical Obstruction of Reverse Transcriptase – One mechanism by which deaminase-independent restriction may occur is the simple obstruction of reverse transcriptase during cDNA synthesis as depicted in **Figure 1-3A** (Bishop *et al.*, 2008; Holmes *et al.*, 2007; Iwatani *et al.*, 2007). Per this model, A3G accumulates on viral genomic RNA and impedes the processivity of reverse transcriptase such that the reaction eventually falters. This mechanism is supported by several lines of evidence. First, A3G has RNA-binding activity (Chelico *et al.*, 2006; Iwatani *et al.*, 2006; Yu *et al.*, 2004b). Second, *in vitro* reverse transcription reactions using A3G or a catalytically inactive A3G mutant inhibit the elongation of reverse transcriptase products (Iwatani *et al.*, 2007). Finally, the length of minus-strand strong-stop cDNA derived from endogenous reverse transcription

reactions in particles – and thus presumably independent of any downstream catabolic pathway – inversely correlates with the amount of A3G or A3G catalytic mutant present (Bishop *et al.*, 2008).

Inhibition of tRNA Priming of Reverse Transcription – Another potential deaminase-independent mechanism of inhibition proposed is the inhibition of tRNA priming during reverse transcription [(Guo *et al.*, 2006; Yang *et al.*, 2007b) and **Figure 1-3A**). Most of these observations were made using overexpressed wildtype A3G or A3F. To extend the observation of decreased reverse transcription to the lack of a deamination requirement, however, these authors employed deletion mutants as their inactive deaminases of choice. Despite the obvious implication that this priming deficiency may be due to A3G's RNA binding activity, either via competition for the primer binding site or perhaps sequestration of the primer itself, this effect has specifically been proposed to be a result of the interaction between A3G and the HIV nucleocapsid protein (Guo *et al.*, 2007). One potential problem with this particular study is that seeing the *in vitro* reverse transcription effect requires A3G:nucleocapsid ratios 10 times greater than those found in naturally restricted virions, but the authors propose that the actual A3G:nucleocapsid ratios present during the critical tRNA annealing step may be more in line with the enhanced relative levels of A3G they used.

Inhibition of Strand Transfer and Integration – A third proposed mechanism of deaminase-independent inhibition by A3G is the inhibition of strand transfer events during reverse transcription (**Figures 1-3B, 1-3D**). One paper describes the inhibition of

both minus- and plus-strand transfers due to a deaminase-independent mechanism of A3G (Li *et al.*, 2007), perhaps caused by the inhibition of RNase H processing via A3G binding to the DNA:RNA heteroduplex, although A3G appears to bind such heteroduplexes relatively weakly (Iwatani *et al.*, 2006; Yu *et al.*, 2004b). Another finds an effect primarily on plus-strand transfer in a deaminase-dependent fashion [(Mbisa *et al.*, 2007) and **Figure 1-2D**). Discrepancies may be due in part to the different mutants used to render A3G catalytically inactive – truncation mutants in the former and point mutants in the latter. Similarly, there have been reports of an effect of A3G on HIV integration, but whether this process is deaminase-dependent (Mbisa *et al.*, 2007) or at least partially deaminase-independent (Luo *et al.*, 2007; Mbisa *et al.*, 2010) is unclear. One interesting mechanistic note contained within one of these reports, however, is the finding that A3G appears to interfere with cleavage of the tRNA primer, creating aberrant ends that may contribute to the downstream inhibition of integration (Mbisa *et al.*, 2007). Exactly how this may occur is unknown, but it echoes speculation on protection of the DNA:RNA heteroduplex (Li *et al.*, 2007).

Discrepancies Among Proposed Deaminase-Independent Mechanisms – Interpretations of deaminase-independent mechanisms of reverse transcription are as complicated as those of the cDNA degradation mechanisms above. One of the primary drawbacks of models focused on early reverse transcription products such as a physical blockage of reverse transcriptase or the inhibition of reverse transcriptase priming is that the inhibition of minus-strand strong stop cDNA is not universally observed under nonpermissive conditions (Li *et al.*, 2004; von Schwedler *et al.*, 1993), and when it is

observed, this effect may be overcome by priming with DNA rather than the natural tRNA^{Lys3} or by removing nucleocapsid from the reaction (Guo *et al.*, 2006; Guo *et al.*, 2007; Li *et al.*, 2007; Yang *et al.*, 2007b). Moreover, even if one finds a defect in early RT products, a specific effect on tRNA priming may not be apparent either (Bishop *et al.*, 2008; Iwatani *et al.*, 2007).

Rather, all authors agree that, whether early reverse transcription is impeded or not, late reverse transcription products are decreased. This may be consistent with those studies proposing effects on strand transfer. Alternatively, this could reflect a cumulative effect of deaminase-independent restriction mechanisms – at multiple steps, in multiple combinations with each other – or any mix of the above. It has even been proposed that A3G may interact directly with reverse transcriptase to mediate the inhibition of proviral cDNA synthesis (Wang *et al.*, 2012) and, similarly, that Vif itself may be a component of reverse transcription complexes that directly facilitates cDNA synthesis (Carr *et al.*, 2008; Carr *et al.*, 2006).

Summary of Deaminase-Independent Restriction Mechanisms – A large body of evidence now supports a variety of deaminase-independent mechanisms of proviral DNA synthesis inhibition. That said, significant discrepancies remain in this literature as in the putative degradation of viral cDNA. The sheer diversity of the deaminase-independent restriction mechanisms described above could merely reflect the fact that RNA binding may be a critical component in all reactions. Alternatively, caution is generally appropriate when interpreting broadly positive results, as they may be explained by a pervasive artifact attributable to specific methods such as transient overexpression of a given protein.

Apropos of pervasive artifacts, several papers now suggest that deaminase-independent mechanisms may be overexpression artifacts [(Browne *et al.*, 2009; Haché *et al.*, 2008; Miyagi *et al.*, 2010; Miyagi *et al.*, 2007; Schumacher *et al.*, 2008) and see **Supplementary Chapter 2**). That is, proponents of deaminase-independent restriction mechanisms universally rely on the transient overexpression of catalytic mutants to make their core functional arguments. When one stably expresses these mutants or titrates back the quantities transiently expressed, however, deaminase-independent restriction is largely absent. As of this writing, there are no published reports of the use of a stably expressed catalytic mutant in combination with a more reductionist approach such as the endogenous reverse transcription reaction. Moreover, despite the potential for more potent deaminase-independent restriction by other factors such as A3F (Holmes *et al.*, 2007), others and we have found that even A3F is incapable of substantial deaminase-independent restriction of HIV when stably expressed [(Miyagi *et al.*, 2010) and see **Supplementary Chapter 2**]. That said, it is possible that the induction of APOBEC3 proteins might theoretically raise expression levels to the range of deaminase-independent relevance (Hultquist *et al.*, 2011; Koning *et al.*, 2009; Refsland *et al.*, 2010).

Mechanism of Vif Neutralization of APOBEC3 Proteins

Many mechanisms are plausible to explain how HIV Vif counteracts the restrictive capacity of APOBEC3 proteins such as A3G. Among published accounts, there is widespread support for the notion that Vif acts to inhibit APOBEC3 encapsidation primarily by marking APOBEC3 proteins for proteasomal degradation and thereby reducing the intracellular levels available for packaging into budding virions [(Conticello *et al.*, 2003; Marin *et al.*, 2003; Mehle *et al.*, 2004b; Sheehy *et al.*, 2003; Stopak *et al.*, 2003; Yu *et al.*, 2003) and **Figure 1-1B**]. Specifically, Vif acts as the substrate specificity adapter within a cullin-RING ubiquitin ligase complex, where it interacts with cullin-5 (CUL5) and elongin C (TCEB1, ELOC) as well as A3G to position the latter for polyubiquitination mediated by the ring-box protein 2 (RBX2/RNF7) bound to the opposite end of the CUL5 scaffold (Conticello *et al.*, 2003; Kao *et al.*, 2003; Marin *et al.*, 2003; Mehle *et al.*, 2004b; Sheehy *et al.*, 2002; Sheehy *et al.*, 2003; Stopak *et al.*, 2003; Yu *et al.*, 2003). Functional stabilization of Vif also requires its interaction with the transcription factor CBFβ in the context of this complex. The specific mechanism involved remains unknown, but it may be as simple as stabilizing Vif steady state expression levels (Hultquist *et al.*, 2012; Jäger *et al.*, 2011; Zhang *et al.*, 2011). By bringing A3G into proximity with the ubiquitin ligase complex, Vif thus facilitates the ubiquitination and eventual degradation of A3G through the proteasome.

A great deal of effort in the field has been expended toward the mapping of specific regions and residues in Vif required for anti-APOBEC3 function. A simplified synthesis of this literature is presented in **Figure 1-4A**. Regions of Vif required for interaction with other members of the E3 ubiquitin ligase complex are well-established.

Chief among these is the single conserved SLQYLA sequence of Vif, the BC box, which is required for interaction between Vif and ELOC [*e.g.* (Mehle *et al.*, 2004a; Yu *et al.*, 2004c)]. Vif may also interact with ELOB, although the conserved PPLP sequence implicated in this activity appears to be multifunctional, as it has also been associated with Vif oligomerization and binding to A3G itself (Bergeron *et al.*, 2010; Donahue *et al.*, 2008; Yang *et al.*, 2003; Yang *et al.*, 2001). Upstream of the BC box but also in the C-terminal half of Vif is an HCCH zinc coordinating motif; these specific amino acids as well as other directly adjacent residues are required for Vif interaction with the CUL5 scaffold but not A3G (Luo *et al.*, 2005; Mehle *et al.*, 2006; Xiao *et al.*, 2007; Yu *et al.*, 2004c).

Functional interaction of Vif with A3G also requires a number of specific residues that, in contrast with those contacting components of the E3 ubiquitin ligase complex, are typically located in the N-terminal half of Vif. While this may appear straightforward, the simplified schematic provided in **Figure 1-4A** and the cognate schematic in **Figure 1-4B** depicting the Vif interaction site and other functionally important residues in A3G gloss over a much more complicated experimental reality. Most residues in Vif have been mutated at this point, many in several independent publications, with sometimes contradictory effects. A summary of this literature is provided in **Table 1-2** in anticipation of **Chapter 5**. To summarize in broad terms, however, the present literature would suggest that the interaction of Vif with different APOBEC3 proteins, for example A3F and A3G, is different. That is, a unique complement of residues is required for interaction with each protein. Some are required for both, but since almost all mutants characterized to date are simple loss of function mutants, it is not clear whether these

regions represent common components of the A3F/A3G-Vif interaction or something less specific. For example, mutation of an amino acid required for overall protein stability might cause loss of Vif expression or gross changes in its structure, either of which may alter the neutralization of APOBEC3 proteins without being part of a normal interaction site. Similarly, changes at residues that interact with some other cellular component required for APOBEC3 neutralization such as CUL5 may be functionally indistinguishable from residues found in a common APOBEC3 interaction site in Vif.

Efforts to identify the Vif interaction site in APOBEC3 proteins have yielded a much more genetically restrained picture. As shown schematically in **Figure 1-4B** and in a full-length model structure of A3G in **Figure 1-4C** (Harjes *et al.*, 2009), there are three prominent residues that, when mutated, can prevent functional interaction of Vif with A3G – DPD128-130, which are in the β4-α4 loop and the very beginning of the α4 helix in A3G (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Lavens *et al.*, 2010; Mangeat *et al.*, 2004; Russell *et al.*, 2009b; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004). Residues with similar phenotypes have also been identified in A3H (D121) and A3F (E289 and E324, see **Chapters 3-5**) and A3D and A3C (Albin *et al.*, 2010b; Smith and Pathak, 2010; Zhen *et al.*, 2010). Unlike the N-terminal Vif-susceptibility residues in A3G, it was somewhat surprising when it was first reported that the Vif interaction region in A3F was not in the N-terminal deaminase domain as in A3G, but rather in the C-terminal deaminase domain (Russell *et al.*, 2009b; Zhang *et al.*, 2008). Given the aforementioned Vif literature, however, this represented still more evidence for the diversity of interactions between Vif and APOBEC3 proteins. As discussed in **Chapters 3** and **5**, however, our own work suggests that the linear separation of these Vif interacting regions

in A3F and A3G is misleading, since the two domains bound are phylogenetically related and, in fact, contain a single conserved structural motif that houses most of the APOBEC3 determinants of Vif susceptibility identified to date (Albin *et al.*, 2010b).

This is a critical point since any attempts to block the interaction of APOBEC3 proteins with Vif would benefit from having one, fixed cellular target at which to bind and prevent Vif interaction, thus theoretically hampering the ability of the virus to evolve resistance to the intervention (Harris and Liddament, 2004). Although no one has yet made public a small-molecule inhibitor of the APOBEC3–Vif interaction, a significant recent development has been the identification of RN-18, a small-molecule antagonist of HIV Vif (Nathans *et al.*, 2008). Interestingly, RN-18 decreases Vif levels and thereby increases A3G (and A3F and A3C but not A3B) levels when these two are present together. This small molecule is also able to inhibit the growth of Vif-proficient HIV in cultured nonpermissive cells. Thus, while further research is necessary to understand the mechanism by which this and other putative inhibitors function, RN-18 has provided proof of principle that the APOBEC3–Vif axis can be targeted by small molecules and, more importantly, that multiple APOBEC3 proteins can be potentiated by a single intervention.

In summary, it is well-established that Vif mediates the degradation of APOBEC3 proteins, thus obviating their potential antiviral functions. It bears mentioning, however, that evidence also exists suggesting that Vif can directly inhibit A3G encapsidation (Kao *et al.*, 2003; Mariani *et al.*, 2003; Sheehy *et al.*, 2003). Such a mechanism need not require degradation as an end point. Furthermore, other reports indicate a Vif-dependent reduction in A3G expression (Kao *et al.*, 2003; Stopak *et al.*, 2003), and it is also possible

33

that Vif may function by directly inhibiting the catalytic activity of A3G (Britan-Rosich *et al.*, 2011; Santa-Marta *et al.*, 2005). These alternative mechanisms are not necessarily mutually exclusive, but an overarching message should be clear: preventing the encapsidation of A3G is crucial for the ability of Vif to preserve HIV infectivity. In fact, reducing the encapsidation of A3G by simply increasing particle production to titrate out the restriction factor is at least one mechanism by which HIV can resist restriction in the absence of Vif function [(Haché *et al.*, 2008) and see **Chapter 6**].

APOBEC3 Proteins *In Vivo*

Beyond restriction and Vif-sensitivity phenotypes, there is one key characteristic that is of some use in distinguishing among APOBEC3 proteins: the sequence context in which they typically mutate DNA cytosines to uracils (**Table 1-1**). A3G preferentially mutates cytosine residues that are preceded by another cytosine, with an even stronger preference for the context 5'-CC<u>C</u>A-3' in minus-strand viral cDNA, where the preferentially targeted cytosine is underlined (Harris *et al.*, 2003a; Mangeat *et al.*, 2003; Yu *et al.*, 2004b; Zhang *et al.*, 2003). In plus-strand reverse-transcribed cDNA, this manifests as 5'-T<u>G</u>GG-3' to 5'-T<u>A</u>GG-3' mutations, or 5'-<u>G</u>G-3' to 5'-<u>A</u>G-3' in the less stringent dinucleotide context. Other APOBEC3 proteins, however, generally display a preference for the mutation of cytosines preceded by another base, most often thymine [*e.g.* (Bishop *et al.*, 2004; Liddament *et al.*, 2004; Wiegand *et al.*, 2004; Zheng *et al.*, 2004)]. This results in 5'-T<u>C</u>-3' to 5'-T<u>U</u>-3' transitions, which manifest on plus-strand viral cDNA as 5'-<u>G</u>A-3' to 5'-<u>A</u>A-3' transitions.

The identification of these different sequence contexts is important for assessing which APOBEC3 proteins act on a given viral substrate in a context where multiple APOBEC3 proteins are expressed, such as *in vivo*. Furthermore, assessing these contexts helps one to appreciate how hypermutation can be detrimental to a virus. Aside from deleterious amino acid or other mutations that may occur in unique sequence contexts, any UGG codon encoding tryptophan is susceptible to the creation of a U<u>A</u>G or perhaps even a U<u>AA</u> stop, particularly in the presence of A3G. Similarly, the AUG start codon may place the wobble base in an APOBEC3-susceptible hotspot depending on the identity of the next base in the sequence, thus making the ablation of initiating

methionines another distinct possibility. Extensive mutation may also disrupt HIV genomic RNA secondary structures, which are important for the regulation of HIV gene expression (Watts *et al.*, 2009).

Even prior to the discovery of APOBEC3 proteins, the tendency for HIV to acquire G-to-A mutations was evident (Janini *et al.*, 2001; Vartanian *et al.*, 1991; Vartanian *et al.*, 1994) – a phenomenon analogous to the original observation of hypermutation in spleen necrosis virus (Pathak and Temin, 1990). These early reports generally noted a preferential sequence context for hypermutation of 5'-<u>G</u>A-3' and, to a lesser extent, 5'-<u>G</u>G-3', suggesting that, given what is now known, multiple APOBEC3 proteins account for these early descriptions of hypermutation (Liddament *et al.*, 2004).

While more recent analyses at times note stronger preferences for one context or another under a given condition (Kieffer *et al.*, 2005; Pace *et al.*, 2006; Piantadosi *et al.*, 2009), a theme uniting recent literature on the *in vivo* hypermutation of HIV is that it frequently occurs in both major dinucleotide contexts. One study indicates, for example, that hypermutated sequences separate roughly into thirds – sequences with hypermutation predominantly in a 5'-<u>G</u>A-3' context, sequences with hypermutation predominantly in a 5'-<u>G</u>G-3' context, and sequences with hypermutation occurring in roughly equal proportions in each context (Land *et al.*, 2008). Another reports the occurrence of 5'-<u>G</u>G-3' context mutations in approximately 95% of hypermutated sequences versus the occurrence of 5'-<u>G</u>A-3' context mutations in approximately 60% of hypermutated sequences (Gandhi *et al.*, 2008). Furthermore, mutations in both contexts frequently occur in the same cloned sequence. For example, 5'-<u>G</u>G-3' hypermutations occur within 95% of predominantly 5'-<u>G</u>A-3' context clones (Gandhi *et al.*, 2008), while another data

set shows a very slight preference for hypermutation in a 5'-GA-3' context with a strong 5'-GG-3' component (Ulenga *et al.*, 2008a).

It is reasonable to conclude, therefore, that multiple APOBEC3 proteins are active *in vivo* and that at least one of these is A3G with its strong 5'-GG-3' context preference. The identities of the APOBEC3 proteins responsible for the reported 5'-GA-3' hypermutations, however, remain unknown. Recent knockout and knockdown studies from our lab indicate that A3F and A3D are likely contributions to 5'-GA-3' context hypermutation in naturally nonpermissive cells; it therefore stands to reason that these are also likely contributing to hypermutation *in vivo* (Refsland *et al.*, 2012). This study did not, however, extensively address the potential for A3H to contribute to 5'-GA-3' context hypermutation in naturally nonpermissive cells, and it stands in contrast with one report implicating A3C in target cell editing of HIV in the 5'-GA-3' context (Bourara *et al.*, 2007). As indicated above, it is also possible that myeloid cell A3A may be a potent, inducible contributor to sequence diversification in the 5'-GA-3' context (Koning *et al.*, 2011; Thielen *et al.*, 2010). In any event, identification of the one true APOBEC3 repertoire may be fundamentally impossible since both expression levels and genotypic differences are likely to contribute to the overall picture in any given patient as discussed below.

Regardless of which APOBEC3 proteins are actively hypermutating HIV, it is clear that their overall effect is significant. Estimates of the frequency of hypermutated sequences *in vivo* range from approximately 7% to 20% with wide variation in the proportions seen in any one individual (Gandhi *et al.*, 2008; Kieffer *et al.*, 2005; Land *et al.*, 2008; Pace *et al.*, 2006; Piantadosi *et al.*, 2009). In fact, these figures almost certainly

underestimate the total proportion of proviral sequences mutated by APOBEC3 proteins. Since a sequence must typically reach a given threshold of G-to-A mutations to be defined as 'hypermutated', methods of quantifying hypermutants fail to account for sequences that have been mutated by APOBEC3 proteins at a lower level. Furthermore, extremely hypermutated sequences may fail to replicate and amplify by PCR while still potentially contributing to the overall mutational load of a viral population through recombination (Mulder *et al.*, 2008). As discussed above, the deamination of cDNA cytosines may also create a substrate for the downstream degradation of the viral genome, which would further diminish the apparent proportion of viral genomes affected by APOBEC3 proteins.

Adaptive potential of G-to-A mutations − A hallmark of HIV is its high mutation rate (Mansky and Temin, 1995), yet this mutation rate likely represents an equilibrium selected by the virus through evolution to optimize host immune evasion without collapsing into error catastrophe (Harris, 2008; Loeb *et al.*, 1999). In this respect, APOBEC3 proteins are of interest not only for their ability to damage viral genomic integrity, but also for their potential ability to change the pace at which the virus may evolve. For example, several authors have written on the theoretical ability of APOBEC3 proteins to introduce mutations into the viral genome that may confer resistance to antiviral drugs, promote immune evasion or even facilitate coreceptor switching (Berkhout and de Ronde, 2004; Bewick *et al.*, 2011; Haché *et al.*, 2006; Pillai *et al.*, 2008).

The levels of mutation characteristic of hypermutation, on the face of it, would seem far too high to be of benefit to HIV. Yet current evidence suggests that there are multiple mechanisms by which HIV may harness the mutagenic power of APOBEC3 proteins without suffering fatal consequences. The first of these centers on the concept of the fitness of a viral population versus that of any individual genome. If a particular viral genome encodes a Vif protein that is hypofunctional in its ability to counteract APOBEC3 proteins, an event that appears to happen at least 20% of the time (Albin *et al.*, 2010a; Simon *et al.*, 2005), processive deamination will introduce a large number of G-to-A mutations into the progeny of that viral genome. Frequently, these may be premature stop codons or mutations ablating the start codon of a given gene, as discussed above, yet recombination can allow the virus to resurrect hypermutated sequences into a functional genomic background (Mulder *et al.*, 2008). Even when hypermutation is at full strength, then, the virus can utilize the expanded sequence space afforded by exposure to APOBEC3 proteins to acquire adaptive mutations.

Even when heavily affected by hypermutation, the virus retains at least two additional ways of averting certain death. One is purely stochastic in that it is possible for a gene to suffer heavy levels of hypermutation but still remain functional (Harris *et al.*, 2003a). A second is so-called purifying selection, a process by which the viral genomes represented at each stage of replication from integrated proviruses through packaged genomes are observed to be progressively less mutated, presumably due to loss of function at various stages of replication – for example, mutations in the viral U3 promoter may select against the transcription of heavily mutated proviruses versus relatively intact counterparts (Russell *et al.*, 2009a).

Given the diverse mechanisms by which HIV may turn hypermutation to its own advantage, it may under certain conditions actually be beneficial for the virus to acquire mutations causing defects in Vif. For example, the presence of the hypofunctional Vif mutant K22H has been correlated with patients experiencing antiretroviral treatment failure, and these patients further display an enrichment for drug resistance mutations in a context consistent with A3G action (Fourati *et al.*, 2010). While this would seemingly suggest that the potentiation of APOBEC3 proteins as a therapeutic strategy may be dangerous, others have found that, despite the potential contribution of A3G and other APOBEC3 proteins to drug resistance mutations, this effect is relatively modest and not necessarily a reason to avoid the therapeutic use of APOBEC3 proteins (Jern *et al.*, 2009).

It may also be possible for HIV to acquire sublethal levels of mutation from APOBEC3 proteins (Sadler *et al.*, 2010), presumably through some residual encapsidation despite the presence of Vif, whether by chance or by viral design. This is echoed by a similar story in which cytotoxic T cell escape mutations consistent with A3G action were found early in infection, implying a role for A3G in providing the sequence diversity on which selection may act from even the earliest stages of infection (Wood *et al.*, 2009). Indeed, even a Vif-proficient virus passaged in the presence of A3G can acquire a drug resistance mutation in a favorable dinucleotide context much more rapidly than the virus itself (Kim *et al.*, 2010). It is thus at least theoretically possible that the virus could benefit from the presence of A3G and other APOBEC3 proteins at some level. Here, it is also important to reiterate the distinction between a viral population and an individual viral genome, as debilitated viruses may suffer mutation rates several times

greater than their wildtype counterparts without necessarily impacting the overall titer of the population [*e.g.* (Haché *et al.*, 2008)].

Clinical assessment of APOBEC3 impacts on HIV infection – While perhaps not entirely appropriate due to the clear *in vitro* effect of APOBEC3 proteins on cDNA accumulation, hypermutation is often taken as a surrogate for APOBEC3 activity in the evaluation of the potential clinical impact of APOBEC3 proteins with an eye toward their use as therapeutics. This approach has yielded a mixed picture for the potential efficacy of APOBEC3 proteins *in vivo* as summarized in **Table 1-3**. A number of publications support a positive correlation between hypermutation levels in patient sequences and reduced viremia and/or increased CD4+ T cell counts (Land *et al.*, 2008; Pace *et al.*, 2006; Vazquez-Perez *et al.*, 2009), but an equal number fail to find any association (Gandhi *et al.*, 2008; Piantadosi *et al.*, 2009; Ulenga *et al.*, 2008a). Aside from the cDNA accumulation confounder, it is possible that this lack of clarity may further suffer from the fact that Vif is presumably largely functional in these patient-derived isolates. Thus, its ongoing suppression of APOBEC3 proteins may be sufficient to lower their apparent activity to a level difficult to assess. Unfortunately, the situation has turned out much the same when assessing APOBEC3 mRNA levels [(Amoedo *et al.*, 2011; Biasin *et al.*, 2007; Cho *et al.*, 2006; Jin *et al.*, 2005; Reddy *et al.*, 2010; Ulenga *et al.*, 2008b; Vazquez-Perez *et al.*, 2009) and reviewed in more detail in (Albin, 2010)]. Among recent positive data, however, an analysis of the effects of interferon induction in HIV/HCV coinfected human patients has shown a modest correlation between decreased viral load and the upregulation of A3F but not A3G (Pillai *et al.*, 2012).

41

The search for APOBEC3 polymorphisms of functional significance has achieved similarly modest success. Polymorphisms in A3H determine stability and restriction activity with the mutations associated with haplotypes II and V being the most relevant to HIV restriction (Cagliani *et al.*, 2011; Harari *et al.*, 2009; OhAinle *et al.*, 2008; Tan *et al.*, 2009; Wang *et al.*, 2011a). The most studied of these two, haplotype II, is present in roughly 50% of individuals of African descent, 10–18% of individuals of European descent and 3–4% of individuals of Asian descent (OhAinle *et al.*, 2008), while a less stable but Vif-resistant haplotype I, which is more common in European and Asian populations, has been correlated with protection from HIV transmission (Cagliani *et al.*, 2011). Another common *APOBEC3* polymorphism is an *APOBEC3B* deletion, which occurs in approximately 20% of the world's population (Kidd *et al.*, 2007). This deletion of *APOBEC3B* is relatively rare in African and European populations but common in East Asian (36.9%), Amerindian (57.7%) and Oceanic (92.9%) populations. Homozygosity for this deletion has previously been associated with enhanced susceptibility to HIV infection and a higher viral set point (An *et al.*, 2009), but an additional study failed to find this effect in a separate cohort (Itaya *et al.*, 2010).

As with other areas of APOBEC3 research, however, the most intensive efforts to date have been directed toward the search for polymorphisms in A3G, and although many have been found, the potential clinical implications are generally unclear [reviewed in (Albin, 2010)]. It has been suggested that one variant, A3G H186R, may be associated with an unfavorable course of infection (An *et al.*, 2004; Reddy *et al.*, 2010), but a smaller study focused on a population where the allele is less prevalent failed to observe these effects (Do *et al.*, 2005). Nevertheless, the H186R variant of A3G is worthy of

further *in vitro* characterization in light of the recent finding that this change results in decreased processivity of the enzyme along ssDNA substrates and consequently decreased deaminase activity (Feng and Chelico, 2011).

It is important to point out, however, that most APOBEC3 proteins have evolved under positive selection, implying that the challenges posed by ancient and presumably divergent pathogens have resulted in the rapid evolution of these restriction factors (OhAinle *et al.*, 2006; Sawyer *et al.*, 2004). While this selective pressure may not have been applied by HIV given its relatively recent arrival in the human population, selection for APOBEC3 proteins better able to restrict other pathogens may have inadvertently selected cross-reactivity for the restriction of HIV. Thus, it is reasonable to predict that, despite the relative dearth of data presently available, variation at the genetic level will account for some effect on HIV *in vivo*. Furthermore, it is possible that ongoing selection by the HIV pandemic itself may result in the expansion of protective *APOBEC3* alleles not yet identified.

CD4$^+$ T cells are the primary reservoir of HIV infection *in vivo*, but relatively little is known about the APOBEC3 repertoire found in these cells. Among activated CD4$^+$ T cells, the Th1 subset of helper T cells expresses greater levels of A3G and A3F, which are further enhanced by the autocrine action of IFNγ (IFNG) (Vetter *et al.*, 2009). Greater expression, in turn, results in greater encapsidation of APOBEC3 proteins in HIV produced from these cells and consequently lower viral infectivity than in the Th2 subset of helper T cells, even for Vif-proficient viruses. In general, most APOBEC3 family proteins with the exception of A3A and, to a lesser extent, A3B appear to be expressed in

primary CD4$^+$ T cells; among T cell subsets, however, A3G is expressed at somewhat higher levels than A3F (Koning *et al.*, 2009).

Studies with primary myeloid lineage cells have also noted a prominent effect of interferon treatment on the ability of APOBEC3 proteins to restrict HIV, linking higher expression of A3G and A3A to the relative resistance of monocytes versus differentiated macrophages to infection (Peng *et al.*, 2007). Another report indicates that A3G and A3F are responsible for preventing HIV from infecting immature dendritic cells, an effect which augments as dendritic cells mature and increase their A3G expression (Pion *et al.*, 2006). Expression of IFNα and, to a lesser extent, IFNγ have also been linked to increased APOBEC3 levels and to increased resistance to HIV infection; this upregulation of APOBEC3 proteins may effectively overwhelm normal Vif function and enable restriction of wildtype HIV in macrophages (Peng *et al.*, 2007; Peng *et al.*, 2006). Of note, IFNα may upregulate all APOBEC3 proteins in macrophages and dendritic cells with the possible exceptions of A3B and A3C, indicating that these are part of the normal interferon response in these cell types, although perhaps not in CD4+ T cells (Koning *et al.*, 2009; Refsland *et al.*, 2010). This upregulation is particularly astounding for A3A (Koning *et al.*, 2009; Peng *et al.*, 2007; Stenglein *et al.*, 2010). Given the common expression of many APOBEC3 proteins in primary cell types and their potential interferon inducibility, it is tempting to speculate that the combination of different APOBEC3 proteins may yield emergent properties not readily apparent in the majority of *in vitro* studies to date, which almost exclusively focus on the effects of single APOBEC3 proteins (See also **Chapter 6**).

Table 1-1: Studies reporting restriction activity, susceptibility to Vif and mutational context preferences of APOBEC3 proteins with respect to HIV.

	Restriction activity against HIV[a]	Susceptibility to Vif[b]	Dinucleotide preference[c]
A3A	Active: (Aguiar *et al.*, 2008; Goila-Gaur *et al.*, 2007; Marin *et al.*, 2008; Rose *et al.*, 2005); note that two of these references involve fusion of A3A to a helper protein	Sensitive: (Marin *et al.*, 2008; Rose *et al.*, 2005)	GA, substantial GG(Aguiar *et al.*, 2008; Stenglein *et al.*, 2010)
	Not active: (Bishop *et al.*, 2004; Bogerd *et al.*, 2008; Bogerd *et al.*, 2006a; Bogerd *et al.*, 2006b; Chen *et al.*, 2006; Gooch and Cullen, 2008; Hultquist *et al.*, 2011; Kinomoto *et al.*, 2007; Wiegand *et al.*, 2004)	Not sensitive: (Hultquist *et al.*, 2011; Marin *et al.*, 2008)	
A3B	Active: (Bishop *et al.*, 2004; Bogerd *et al.*, 2007; Bogerd *et al.*, 2006a; Bogerd *et al.*, 2006b; Doehle *et al.*, 2005a; Hakata and Landau, 2006; Hultquist *et al.*, 2011; Kinomoto *et al.*, 2007; Rose *et al.*, 2005; Yu *et al.*, 2004a)	Sensitive: no published reports	GA: (Bishop *et al.*, 2004; Doehle *et al.*, 2005a)
	Not active: (Haché *et al.*, 2008; Hultquist *et al.*, 2011; Refsland *et al.*, 2012), note the distinction between spreading infection and 293T overexpression systems	Not sensitive: (Bishop *et al.*, 2004; Doehle *et al.*, 2005a; Hultquist *et al.*, 2011; Rose *et al.*, 2005; Yu *et al.*, 2004a)	
A3C	Active: (Bishop *et al.*, 2004; Bogerd *et al.*, 2006a; Bogerd *et al.*, 2006b; Kinomoto *et al.*, 2007; Langlois *et al.*, 2005; Marin *et al.*, 2008; Yu *et al.*, 2004a)	Sensitive: (Chen *et al.*, 2009; Hultquist *et al.*, 2011; Langlois *et al.*, 2005; Marin *et al.*, 2008; Pery *et al.*, 2009; Zhang *et al.*, 2008)	GA, substantial GG [nonspecific compared with others (Langlois *et al.*, 2005; Yu *et al.*, 2004a)]
	Not active: (Doehle *et al.*, 2005a; Hultquist *et al.*, 2011; Refsland *et al.*, 2012; Rose *et al.*, 2005; Wiegand *et al.*, 2004; Zheng *et al.*, 2004)	Not sensitive: (Bishop *et al.*, 2004; Yu *et al.*, 2004a)	
A3D	Active: (Chen *et al.*, 2009; Dang *et al.*, 2008a; Dang *et al.*, 2006; Hultquist *et al.*, 2011; Kinomoto *et al.*, 2007; Refsland *et al.*, 2012; Zhang *et al.*, 2008)	Sensitive: (Chen *et al.*, 2009; Dang *et al.*, 2008a; Dang *et al.*, 2006; Hultquist *et al.*, 2011; Zhang *et al.*, 2008)	GA, substantial GT(Dang *et al.*, 2006; Refsland *et al.*, 2012)
	Not active: (Haché *et al.*, 2008)	Not sensitive: no published reports	

	Restriction activity against HIV[a]	Susceptibility to Vif[b]	Dinucleotide preference[c]
A3F	Active: (Bishop et al., 2004; Liddament et al., 2004; Wiegand et al., 2004; Zheng et al., 2004) and many others; restriction by A3F is generally well accepted, and A3F is often used as a control with A3G	Sensitive: (Bishop et al., 2004; Liddament et al., 2004; Wiegand et al., 2004; Zheng et al., 2004)	GA: (Bishop et al., 2004; Liddament et al., 2004; Wiegand et al., 2004; Zheng et al., 2004)
	Not active: (Miyagi et al., 2010; Mulder et al., 2010)	Not sensitive: no published reports, although A3F may be generally less responsive to Vif than A3G [e.g. (Liddament et al., 2004)]	
A3G	Active: (Harris et al., 2003a; Mangeat et al., 2003; Sheehy et al., 2002; Zhang et al., 2003) plus countless reports since	Sensitive: (Conticello et al., 2003; Mehle et al., 2004a; Mehle et al., 2004b; Sheehy et al., 2002; Sheehy et al., 2003; Stopak et al., 2003; Yu et al., 2003) plus countless reports since	GG: (Harris et al., 2003a; Liddament et al., 2004; Mangeat et al., 2003; Yu et al., 2004b; Zhang et al., 2003)
	Not active: no published reports	Not sensitive: no published reports	
A3H	Active: (Dang et al., 2008a; Harari et al., 2009; Hultquist et al., 2011; Li et al., 2010; OhAinle et al., 2008; Refsland et al., 2012; Tan et al., 2009; Wang et al., 2011a)	Sensitive: (OhAinle et al., 2008; Tan et al., 2009) Partially sensitive(Hultquist et al., 2011; Li et al., 2010; Refsland et al., 2012)	GA(Harari et al., 2009)
	Not active: (Dang et al., 2006; Kinomoto et al., 2007; OhAinle et al., 2006); multiple alleles and splice variants exist in the human population, not all of which are active or equally impacted by Vif, as briefly discussed in main text	Not sensitive: (Dang et al., 2008a; Harari et al., 2009; Wang et al., 2011a)	

[a]Single-cycle activity here is defined as the suppression of infectivity relative to negative restriction controls of at least 50%. Spreading infection activity is defined as the substantial suppression of viral spread for the duration of the experiment shown. Some experiments involved the modification of APOBEC3 proteins by, for example, fusion to a helper protein. Where this occurs, restriction is counted as if the protein were wildtype.
[b]Any data concerning the Vif sensitivity of modified APOBEC3 proteins as described above are disregarded in this table.
[c]References for the dinucleotide preferences of APOBEC3 mutational activity are preferentially chosen from papers with data on mutation of the HIV genome. Studies on other substrates generally agree with those listed here, but variation among papers may occur.

Table 1-2: Anti-APOBEC3 loss of function Vif mutants.

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
E2	98%				A	A	(Yamashita *et al.*, 2008)
W5	99%	A	A		A	A	(Simon *et al.*, 1999; Tian *et al.*, 2006)
Q6	99%				A	A	(Simon *et al.*, 1999)
V7	95%				A	A	(Russell and Pathak, 2007)
M8	81%				A	A	(Russell and Pathak, 2007)
I9	97%				A	A	(Russell and Pathak, 2007)
V10	99%				A	A	(Russell and Pathak, 2007)
W11	99%	A	A, R		A, R, L	L	(Mulder *et al.*, 2010; Russell and Pathak, 2007; Simon *et al.*, 2005; Tian *et al.*, 2006; Yamashita *et al.*, 2008)
Q12	99%		A		A	A	(Pery *et al.*, 2009; Russell and Pathak, 2007; Simon *et al.*, 1999)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
V13	98%		A		A	A	(Pery *et al.*, 2009; Russell and Pathak, 2007; Simon *et al.*, 1999)
D14	99%		A	A, S	A		(He *et al.*, 2008; Pery *et al.*, 2009; Russell and Pathak, 2007; Zhang *et al.*, 2008)
R15	99%		A	A, E	A		(He *et al.*, 2008; Pery *et al.*, 2009; Russell and Pathak, 2007; Zhang *et al.*, 2008)
M16	99%		A		A	A	(Russell and Pathak, 2007; Simon *et al.*, 1999)
R17	88%		A	Q	A	A	(Pery *et al.*, 2009; Russell and Pathak, 2007; Simon *et al.*, 1999)
I18	99%				A	A	(Russell and Pathak, 2007; Simon *et al.*, 1999)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
W21	99%	A	A				(Nagao *et al.*, 2010; Tian *et al.*, 2006; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010)
K22	46%	E, H, D	E		A, R	A, D, R, E	(Chen *et al.*, 2009; Dang *et al.*, 2009; Fourati *et al.*, 2010; Simon *et al.*, 2005; Zhang *et al.*, 2008)
S23	98%	Del	Del		A	A	(Chen *et al.*, 2009; Dang *et al.*, 2009; Simon *et al.*, 1999)
L24	97%	Del, S	Del, S		A	A	(Chen *et al.*, 2009; Dang *et al.*, 2009; Nagao *et al.*, 2010; Simon *et al.*, 1999; Yamashita *et al.*, 2008)
V25	98%	Del	Del, S		S		(Chen *et al.*, 2009; Dang *et al.*, 2009)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
K26	99%	A, D, R, Y, Q				A, D, R, Y, Q	(Albin *et al.*, 2010a; Chen *et al.*, 2009; Dang *et al.*, 2009)
H27	81%				Q	Q	(Albin *et al.*, 2010a)
H28	99%				A	A	(Nagao *et al.*, 2010; Yamashita *et al.*, 2008)
M29	85%	A	A				(Simon *et al.*, 1999)
Y30	78%	A				A	(Chen *et al.*, 2009)
S32	99%	A, P	A, P				(Nagao *et al.*, 2010; Simon *et al.*, 2005; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010)
R33	65% K				A	A	(Simon *et al.*, 1999)
K34	83%	D	D		A, R	A, R	(Dang *et al.*, 2009; Simon *et al.*, 1999)
R36	61% K				A		(Chen *et al.*, 2009)
D37	59% G	A	A				(Russell and Pathak, 2007)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
W38	99%	A	A				(Nagao *et al.*, 2010; Russell and Pathak, 2007; Tian *et al.*, 2006; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010)
F39	51%				A	A	(Russell and Pathak, 2007)
Y40	99%	A, H				A, H	(Chen *et al.*, 2009; He *et al.*, 2008; Nagao *et al.*, 2010; Russell and Pathak, 2007; Simon *et al.*, 2005; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010)
R41	83%	A				A	(Chen *et al.*, 2009; He *et al.*, 2008; Russell and Pathak, 2007; Zhang *et al.*, 2008)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
H42	98%	A, N				A, N	(Chen *et al.*, 2009; He *et al.*, 2008; Pery *et al.*, 2009; Russell and Pathak, 2007; Zhang *et al.*, 2008)
H43	99%	A, N				A, N	(Nagao *et al.*, 2010; Pery *et al.*, 2009; Russell and Pathak, 2007; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010)
Y44	86%	A				A	(Russell and Pathak, 2007)
E45	89%	G	G		A	A	(Russell and Pathak, 2007; Simon *et al.*, 2005)
S46	88%				A	A	(Russell and Pathak, 2007)
T47	38% R				A	A	(Russell and Pathak, 2007)
N48	76% H	A				A	(Russell and Pathak, 2007)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
P49	99%	A	A				(Russell and Pathak, 2007)
K50	49%	A	A				(Russell and Pathak, 2007)
I51	55%				A	A	(Russell and Pathak, 2007)
S52	97%				A	A	(Russell and Pathak, 2007)
S53	99%				A	A	(Simon *et al.*, 1999; Yamashita *et al.*, 2008)
E54	98%				A	A	(Simon *et al.*, 1999)
V55	97%				S, A	A	(He *et al.*, 2008; Pery *et al.*, 2009)
I57	98%				S, A	A	(He *et al.*, 2008; Pery *et al.*, 2009)
P58	99%				A	A	(He *et al.*, 2008; Simon *et al.*, 1999; Yamashita *et al.*, 2008)
L59	94%				S, A	S, A	(He *et al.*, 2008; Simon *et al.*, 1999)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
L64	94%	S	S		A	A	(He *et al.*, 2008; Pery *et al.*, 2009)
I66	89%	S	S		A	A	(He *et al.*, 2008; Pery *et al.*, 2009)
Y69	99%	A	A	A	F		(He *et al.*, 2008; Nagao *et al.*, 2010; Pery *et al.*, 2009; Yamashita *et al.*, 2008)
W70	99%	A	A		A	A	(He *et al.*, 2008; Pery *et al.*, 2009; Tian *et al.*, 2006)
L72	98%	S, A	S, A	A			(He *et al.*, 2008; Pery *et al.*, 2009)
H73	69%				A	A	(Simon *et al.*, 1999)
T74	91%				A	A	(He *et al.*, 2008; Simon *et al.*, 1999)
E76	99%		A		A	A	(He *et al.*, 2008; Nagao *et al.*, 2010; Yamashita *et al.*, 2008)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
R77	83%				A	A	(He *et al.*, 2008)
W79	99%		A	A	A		(He *et al.*, 2008; Tian *et al.*, 2006; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010; Zhang *et al.*, 2008)
H80	91%				A	A	(Simon *et al.*, 1999)
L81	98%		S		S, A	A	(Dang *et al.*, 2010a; Simon *et al.*, 1999)
G82	99%		D		D		(Dang *et al.*, 2010a)
G84	99%	D	D, A		A		(Dang *et al.*, 2010a)
V85	92%				S	S	(Dang *et al.*, 2010a)
S86	99%				A	A	(Dang *et al.*, 2010a; Simon *et al.*, 1999; Yamashita *et al.*, 2008)
I87	96%				A	A	(Dang *et al.*, 2010a; Simon *et al.*, 1999)
E88	98%				A	A	(Dang *et al.*, 2010a)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
W89	99%	A	A		A	A	(Dang *et al.*, 2010a; Tian *et al.*, 2006)
R90	93%				D, A	D, A	(Dang *et al.*, 2010a; Simon *et al.*, 1999)
K91	44%				D, A	D, A	(Dang *et al.*, 2010b; Simon *et al.*, 1999)
R92	48%				D, A	D, A	(Dang *et al.*, 2010b; Simon *et al.*, 1999)
R93	74%				D	D	(Dang *et al.*, 2010b)
Y94	98%				A	A	(Dang *et al.*, 2010b; Yamashita *et al.*, 2008)
S95	76%				A	A	(Dang *et al.*, 2010b)
T96	99%	D, R	D, R		A	A	(Dang *et al.*, 2010b; Yamashita *et al.*, 2008)
Q97	99%		R		A, R	A	(Dang *et al.*, 2010b; Simon *et al.*, 1999)
V98	72%				S, A	S, A	(Dang *et al.*, 2010b; Simon *et al.*, 1999)
D99	93%				K	K	(Dang *et al.*, 2010b)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
P100	99%				S, A	S, A	(Dang *et al.*, 2010b)
D101	42%				R	R	(Dang *et al.*, 2010b)
L102	91%				S	S	(Dang *et al.*, 2010b)
A103	99%	D, R, Y	D, R, Y		S, V	S, V	(Dang *et al.*, 2010b)
D104	99%	A, R, Y	A, R, Y		S	A, S	(Dang *et al.*, 2010b; Nagao *et al.*, 2010)
Q105	88%				A	A	(Dang *et al.*, 2010b)
L106	99%				S	S	(Dang *et al.*, 2010b)
I107	99%	R	R		A	A	(Dang *et al.*, 2010b)
H108*	99%	A	A				(Nagao *et al.*, 2010; Yamashita *et al.*, 2010)
Y111	97%				A	A	(Yamashita *et al.*, 2008)
C114*	99%	A, S	A, S				(Nagao *et al.*, 2010; Simon *et al.*, 2005)
F115	99%	S	S				(Simon *et al.*, 2005)
S118	96%				A	A	(Yamashita *et al.*, 2008)
R121	98%				A	A	(Simon *et al.*, 1999)
N122	49%				A	A	(Simon *et al.*, 1999)
T123	97%				A	A	(Simon *et al.*, 1999)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
R127	53% H				A	A	(Simon *et al.*, 1999)
I128	49% I				A	A	(Simon *et al.*, 1999)
C133*	99%	A, S	A, S				(Nagao *et al.*, 2010; Simon *et al.*, 1999)
G138	99%	R	R				(Simon *et al.*, 2005)
H139*	99%	A	A				(Nagao *et al.*, 2010)
G143	98%	R	R				(Simon *et al.*, 2005)
Y147	98%				A	A	(Simon *et al.*, 1999)
L148	99%				A	A	(Simon *et al.*, 1999; Yamashita *et al.*, 2008)
L150		P	P				(Simon *et al.*, 2005)
P156	93%				A	A	(Simon *et al.*, 1999)
K157	78%				A	A	(Simon *et al.*, 1999)
Q158	65% K				A	A	(Simon *et al.*, 1999)
I159	39%				A	A	(Simon *et al.*, 1999)
K160	86%				A	A	(Simon *et al.*, 1999)
P161	98%	A	A		A	A	(Simon *et al.*, 1999)
P162	98%	A	A		A	A	(Simon *et al.*, 1999; Yamashita *et al.*, 2008)

Residue	Conserved	Inactive vs. A3G	Inactive vs. A3F	Inactive vs. A3C and A3D	Active vs. A3G	Active vs. A3F	Refs
L163	98%	A	A		A	A	(Simon *et al.*, 1999)
S165	99%				A	A	(Yamashita *et al.*, 2008)
V166	92%				A	A	(Simon *et al.*, 1999)
L169	99%				A	A	(Simon *et al.*, 1999)
T170	69%				A	A	(Simon *et al.*, 1999)
E171	99%		A		A	A	(Dang *et al.*, 2010a; Yamashita *et al.*, 2008)
D172	98%		A		A		(Dang *et al.*, 2010a)
R173	97%		A		A		(Dang *et al.*, 2010a)
W174	99%		A		A		(Dang *et al.*, 2010a)
N175	98%				A	A	(Dang *et al.*, 2010a)
T180	86%				A	A	(Simon *et al.*, 1999)
K181	60%				A	A	(Simon *et al.*, 1999)
M189	92%				A	A	(Simon *et al.*, 1999)
W190	84%				A	A	(Simon *et al.*, 1999)

Columns indicate the wildtype amino acid at each residue in a given study, its conservation among HIV isolates, loss of function mutations against A3G, A3F and A3D and A3C, aphenotypic mutations against A3G and A3F and the references reporting said mutants. Mutants are considered inactive if they demonstrate a loss of function of approximately 50% or greater relative to wildtype, with some deference given to the authors' own interpretations.

Known zinc coordinating residues that affect interaction with CUL5 are indicated with an asterisk. References focusing on these and other residues required for interaction with other components of the E3 ligase complex are cited in the main text. Citations in this table derive from papers specifically seeking loss of function with respect to APOBEC3 neutralization.

Where two or more mutations are made simultaneously, a situation that typically occurs among contiguous residues, each point mutation is considered as though it were the only mutation present in the construct. The exception to this is the PPLP motif in (Simon *et al.*, 1999), where the phenotypes of the individual amino acid changes in isolation are included in addition to the phenotypes associated with combination mutants.

In (Yamashita *et al.*, 2008) and in (Simon *et al.*, 1999), growth in H9 cells is taken as intact anti-A3F and anti-A3G function. Loss of function in this cell line is not considered due to inability to distinguish A3F from A3G or other APOBEC3 proteins.

Conservation derived from the 2011 Los Alamos National Laboratory HIV Sequence Database curated alignment for HIV subtypes ("All M Group A-K + Recombinants").

Table 1-3. Summary of linkage studies between clinical indicators and hypermutation, APOBEC3 expression or APOBEC3 polymorphism.

Study	Cohort description	Relationships reported
Association with hypermutation		
(Pace *et al.*, 2006)	136 Australians, largely male and Caucasian; all treatment-naïve	Reduced viremia attributable to hypermutation
(Gandhi *et al.*, 2008)	8 elite suppressors and 9 patients on highly active antiretroviral therapy	Hypermutation comparable between the two groups with the exception of one elite suppressor who had significantly elevated hypermutation levels
(Land *et al.*, 2008)	215 Kenyan female sex workers and 25 Kenyan women of similar socioeconomic status who were not commercial sex workers; all treatment-naïve	Increased CD4$^+$ cell counts in 17 patients with significant hypermutation Positive correlation between adenine content and CD4$^+$ cell counts among 208 patients from whom counts were available
(Ulenga *et al.*, 2008a)	29 Senegalese female sex workers; all treatment-naïve	No correlation between hypermutation and viremia
(Vazquez-Perez *et al.*, 2009)	45 HIV-infected, 37 HIV-exposed seronegative, and 26 healthy control patients; all treatment-naïve	Increased hypermutation in patients with low viral load
(Piantadosi *et al.*, 2009)	28 Kenyan women; all treatment-naïve	No correlation between percentage of hypermutated sequences and either viral load in chronic infection or CD4$^+$ cell counts approximately 6 weeks postinfection No difference for either indicator between patients with and without hypermutated sequences
(Amoedo *et al.*, 2011)	17 Brazilian, perinatally-infected pediatric AIDS patients, most treatment-experienced	No correlation between hypermutation levels and progressor/nonprogressor status

Study	Cohort description	Relationships reported
Association with APOBEC3 expression		
(Jin *et al.*, 2005)	6 uninfected individuals and 25 infected individuals consisting of 8 long-term nonprogressors and 17 progressors; all treatment-naïve	Inverse correlation between A3G expression and viral load Positive correlation between A3G expression and CD4$^+$ cell counts among the 25 infected individuals Higher A3G mRNA levels in long-term nonprogressors and uninfected individuals relative to progressors
(Cho *et al.*, 2006)	92 infected individuals (~2/3 African American and ~1/3 Caucasian) and 19 uninfected individuals (~2/3 Caucasian); infected individuals off treatment for at least three months	Lower A3F and A3G mRNA levels in infected versus uninfected individuals No correlation between A3F or A3G mRNA levels and viral load or CD4$^+$ cell counts
(Biasin *et al.*, 2007)	30 each of HIV-exposed seronegative individuals, HIV-infected individuals and uninfected individuals; all infected individuals on antiretroviral therapy	Greater A3G protein levels in exposed seronegative individuals compared with healthy controls or infected individuals Cells from exposed seronegative individuals relatively resistant to *ex vivo* infection compared with healthy controls
(Ulenga *et al.*, 2008b)	Senegalese female sex workers – 16 with a low viral set point and 14 with a high viral set point; all treatment-naïve	Inverse correlation between A3G or A3F mRNA levels after infection and viral set point
(Vazquez-Perez *et al.*, 2009)	45 infected, 37 HIV-exposed seronegative and 26 healthy control patients; all treatment-naïve	Positive correlation between A3G mRNA and CD4$^+$ cell counts in infected patients Negative correlation between A3G mRNA expression and viremia Higher A3G mRNA expression in exposed seronegative individuals compared with healthy controls
(Pillai *et al.*, 2012)	19 HIV/HCV coinfected patients receiving IFNα/ribavirin therapy	Inverse correlation between interferon-mediated upregulation of A3F but not A3G and viral load
(Amoedo *et al.*, 2011)	17 Brazilian, perinatally-infected pediatric AIDS patients, most treatment-experienced	No correlation between A3F/G mRNA expression levels and progressor/nonprogressor status

Study	Cohort description	Relationships reported
Association with APOBEC3 polymorphisms		
(An *et al.*, 2004)	Total of 3073 patients in six cohorts: 1481 European Americans and 949 African Americans broken into groups of seroconverters, seroprevalents and seronegatives among patients in the USA; 643 seroprevalent Swiss individuals	7 polymorphisms identified in a subset of 92 African American and 92 European American patients H186R also identified in Do *et al.* (Ref. 105) more common in African Americans and correlates with loss of CD4$^+$ cells and progression to AIDS in this group
(Do *et al.*, 2005)	245 HIV-infected slow progressors, 82 HIV-infected rapid progressors and 446 healthy subjects; all French Caucasians	None of 29 polymorphisms identified in A3G associated with rate of disease progression
(Pace *et al.*, 2006)	136 Australians, largely male and Caucasian; all treatment-naive	22 polymorphisms identified No association between two closely examined polymorphisms and hypermutation levels
(Valcke *et al.*, 2006)	122 Canadian Caucasian males	6 polymorphisms identified C40693T intronic change associated with increased risk of infection
(An *et al.*, 2009)	4216 infected individuals from five longitudinal cohorts based in the USA	Homozygous A3B deletion associated with increased risk of infection, faster progression to AIDS and increased viral set point No association detected in hemizygous individuals
(De Maio *et al.*, 2011)	Pediatric HIV infected cohort	No association between H186R or C40693T A3G SNPs and progression to AIDS
(Cagliani *et al.*, 2011)	70 Italian HIV-exposed seronegative partners in serodiscordant relationships	Association of A3H Haplotype I with preventing infection
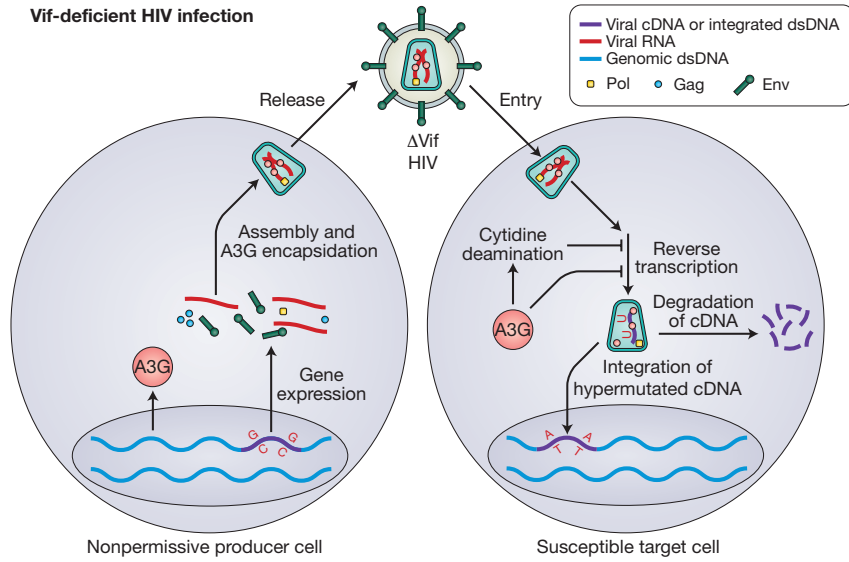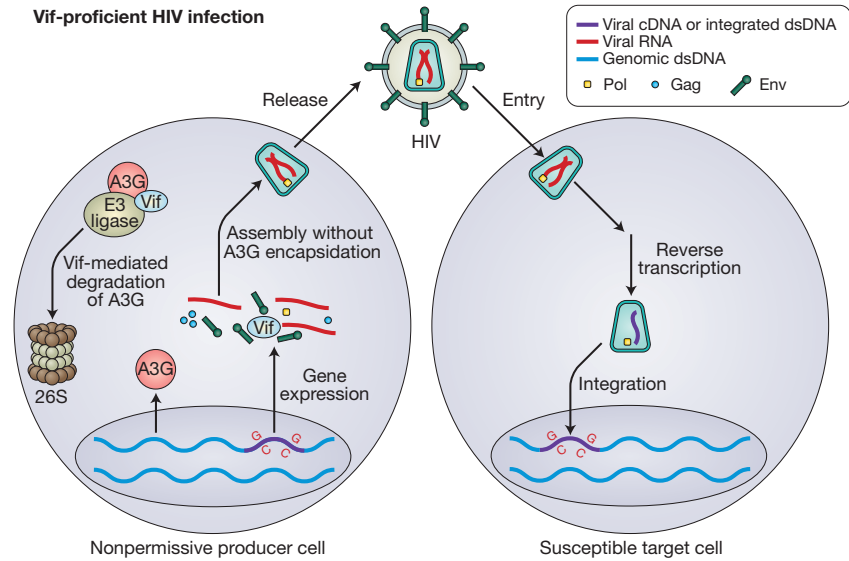
Figure 1-1: APOBEC3 and Vif function.

Figure 1-1: APOBEC3 and Vif function. (A) Vif-deficient HIV infection using the prototypical APOBEC3 protein A3G as an example. In a 'nonpermissive producer cell' – that is, one that expresses the DNA cytosine deaminase A3G and/or other restrictive APOBEC3 proteins and does not support replication of Vif-deficient HIV – gene expression of Vif-deficient HIV proceeds normally but, during assembly, A3G (pink circles) is encapsidated along with normal virion components. Virions with encapsidated A3G bud from a producer cell and enter a target cell normally. Reverse transcription, however, is blocked by the presence of A3G, either by the direct inhibition of cDNA synthesis or the degradation of uracil-containing cDNA, arising from cytosine deamination. Uracil-containing cDNA templates A on the viral plus-strand, resulting in G-to-A hypermutation. A direct A3G-mediated block to integration has also been proposed (not shown). **(B)** Vif-proficient HIV infection. To prevent encapsidation, Vif expressed from a Vif-proficient virus links A3G to an E3 ligase complex, which results in the ubiquitination of A3G and its eventual degradation in the proteasome. Alternatively, Vif may directly inhibit encapsidation, lower A3G expression and/or inhibit A3G catalytic activity (not shown). Subsequent release, entry, reverse transcription and integration are shown proceeding normally.
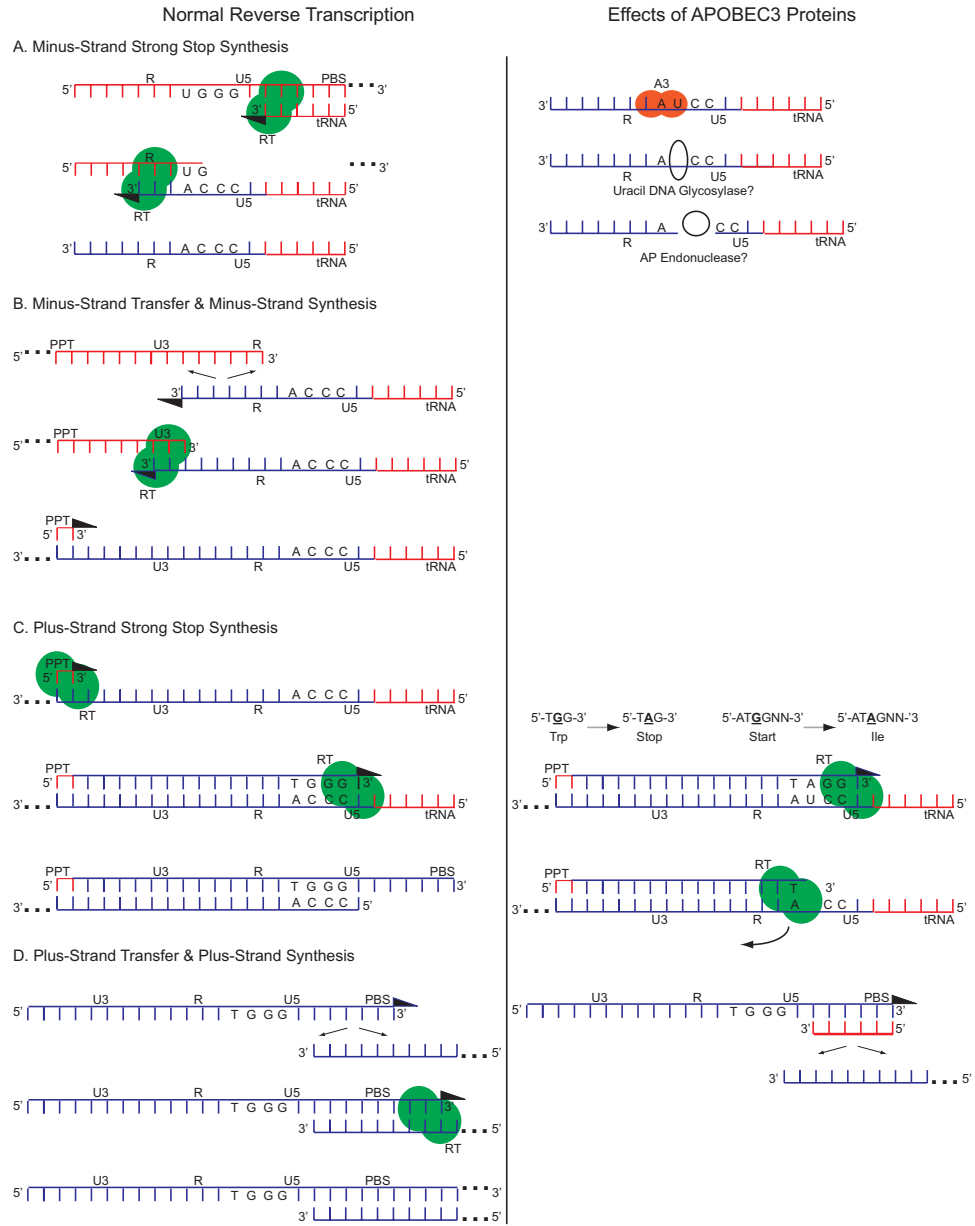
Figure 1-2: Steps of reverse transcription impacted by deaminase-dependent mechanisms of APOBEC3 restriction.

Figure 1-2: Steps of reverse transcription impacted by deaminase-dependent mechanisms of APOBEC3 restriction. Normal reverse transcription is depicted in the left column, while points at which APOBEC3 proteins may impede reverse transcription by deaminase-dependent mechanisms are depicted at the right. References are distributed throughout the main text. **(A)** APOBEC3 deamination of a target cytosine to uracil creates a substrate on which a uracil DNA glycosylase may act (blank oval), creating an abasic site that may be further processed by an unspecified AP endonuclease to degrade viral cDNA (blank circle). **(B)** Steps unaffected per published mechanisms of APOBEC3 deaminase-dependent restriction. **(C)** Persistence of uracil in the viral minus strand is a mutagenic event that, among many possibilities, may introduce premature stop codons or ablate initiating methionine codons. Alternatively, persistence of an abasic site subsequent to APOBEC3-dependent deamination and processing by a uracil DNA glycosylase enzyme may impede synthesis of the plus-strand or, alternatively, result in the insertion of adenine across from the abasic site, yielding the same mutagenic event as minus-strand uracil persistence. **(D)** APOBEC3 deaminase-dependent aberrant RNase H processing of the tRNA primer may prevent liberation of the viral plus-strand for transfer to the 5' end of the genome and subsequent priming of additional plus-strand synthesis.
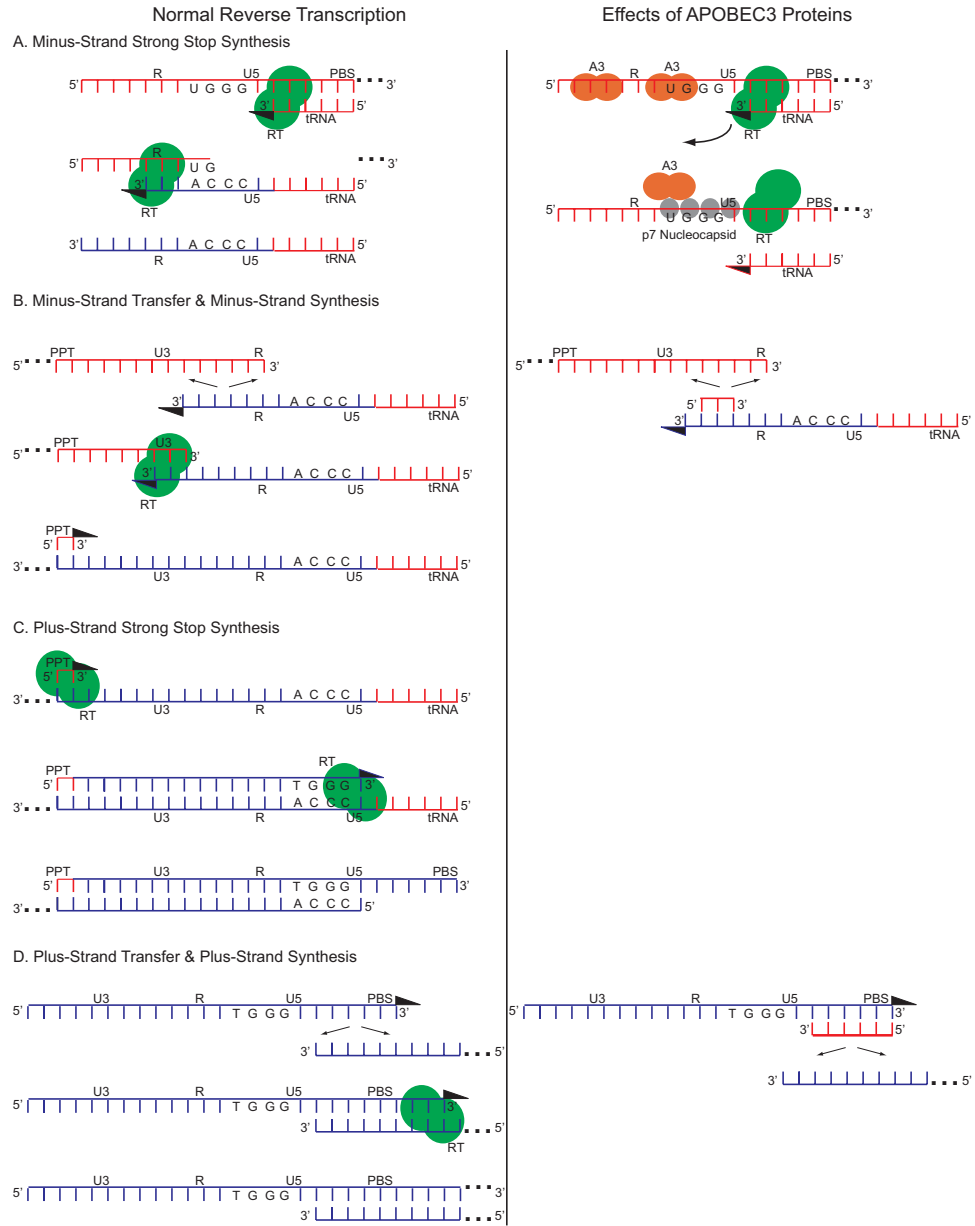
Figure 1-3: Steps of reverse transcription impacted by deaminase-independent mechanisms of APOBEC3 restriction.

Figure 1-3: Steps of reverse transcription impacted by deaminase-independent mechanisms of APOBEC3 restriction. Normal reverse transcription is depicted in the left column, while points at which APOBEC3 proteins may impede reverse transcription by deaminase-independent mechanisms are depicted at the right. References are distributed throughout the main text. **(A)** Binding of APOBEC3 proteins to genomic RNA may impede the processivity of reverse transcriptase, preventing the completion of cDNA synthesis. Alternatively, the reverse transcription reaction may fail to prime in a manner dependent on APOBEC3 interaction with the viral nucleocapsid protein (gray circles). **(B)** APOBEC3 deaminase-independent aberrant RNase H processing of the template genomic RNA may fail to liberate the minus-strand strong stop cDNA to transfer to the 3' end of the genome and prime additional minus-strand synthesis. **(C)** Steps unaffected per published mechanisms of APOBEC3 deaminase-independent restriction. **(D)** APOBEC3-dependent aberrant RNase H processing of the tRNA primer may fail to liberate the plus-strand for transfer to the 5' end of the genome and subsequent priming of additional plus-strand synthesis.

A HIV-1 Vif

Interaction with
A3F, A3D and A3C

Interaction
with A3G

Interaction
with A3G

Interaction with
A3F and A3G

Interaction with
A3F, A3DE
and A3C

HCCH zinc-binding domain:
interaction with CUL5

BC box: interaction
with Elongin C

1 ^{14}DRMR^{17} ^{40}YRHHY^{44} ^{55}Amino acids^{72} ^{108}H ^{114}C ^{133}C ^{139}H 192

^{23}SLV^{25}
^{22}K ^{26}K

^{74}TGERDW^{79}

^{144}SLQYLA^{149}

B Human APOBEC3G

Encapsidation Interaction
with Vif

N-terminal
pseudoactive site

C-terminal
active site

1 ^{65}HxE^{67} ^{96}PCxxC^{100} ^{126}FWDPDYQ^{132} ^{257}HxE^{259} ^{287}PCxxC^{291} 384

^{124}YYFW^{127}

C

C

Active site

Vif binding

D128

Encapsidation

Active site

Pseudoactive site

180°

N

N

Figure 1-4: Important domains in Vif and APOBEC3G.

70

Figure 1-4: Important domains in Vif and APOBEC3G. While many papers have identified residues critical to Vif function, only putative continuous interaction domains in Vif **(A)** and APOBEC3G (A3G) **(B)** are depicted here for the sake of clarity. Exceptions are made for two important lysines in Vif (K22 and K26) as well as the C-terminal zinc-binding domain of Vif. Vif and A3G are internally to scale, but Vif is depicted at twice its actual size relative to A3G. HIV Vif residues shown are those found in the HXB2 reference sequence (accession number K03455). A3G corresponds to reference sequence NP_068594. **(C)** Important domains of A3G identified by genetic analysis are highlighted on a full-length model structure of APOBEC3G (A3G) (Harjes *et al.*, 2009) with coloration carried out in Pymol. See main text for references and discussion. Side chains are shown for key regions, and the right panel represents a rotation of 180º about the *y*-axis. Red spheres, active-site zinc; green side chains (^{124}YYFW^{127}), encapsidation determinant (overlaps 126–127 start of Vif-binding region ^{126}FWDPDYQ^{132}); red sidechain (D128), a key residue for Vif sensitivity; blue sidechains (^{129}PDYQ^{132}), the remainder of the putative Vif-binding site; yellow side chains (H65 & E67 and H257 & E259), the HxE of the pseudocatalytic and catalytic N- and C-terminal domains, respectively; orange side chains (^{96}PC^{97} & C100 and ^{287}PC^{288} & C291), the PCxxC of the active and pseudoactive N- and C-terminal domains, respectively.

Chapter 2: Long-term Restriction by APOBEC3F Selects Human

Immunodeficiency Virus Type 1 Variants with Restored Vif Function

Foreword

This chapter is reproduced with permission from the following publication:

> Albin JS, Haché G, Hultquist JF, Brown WL and Harris RS. "Long-term restriction by APOBEC3F selects human immunodeficiency virus type 1 variants with restored Vif function." *Journal of Virology*. 2010; 84 (19): 10209-10219.

The work described in this chapter is the continuation of a project initiated by a previous student, Guylaine Haché. She described the evolution of Vif-deficient isolates that bypass A3G by repairing an aberrant upstream open reading frame in the parent HIV$_{IIIB}$ strain, leading to enhanced viral production and thereby "tolerance" of the presence of A3G despite the persistent absence of Vif (Haché *et al.*, 2009; Haché *et al.*, 2008).

One observation to come of these studies was that repair of this aberrant upstream open reading frame was insufficient to permit growth in cell lines stably transfected with restrictive levels of A3F. We therefore set out to determine how viruses might bypass A3F by carrying out similar virus evolution studies starting from viral templates with a repaired upstream open reading frame and premature *vpr* stop codons. The function of Vpr deficiency in these and related studies remains obscure, but it is a common observation among viruses passaged in culture and may represent a culture-adaptive feature [*e.g.* (Nakaya *et al.*, 1994) and see **Chapters 5** and **6**]. For example, inactivation of Vpr and its associated cell cycle arrest activity may keep infected cells alive longer to produce more virions. In any event, the Vpr status of the viruses described in this manuscript is, as we will show, irrelevant to their ability to neutralize A3F.

Summary

Tandem stop mutations K26X and H27X in human immunodeficiency virus type 1 (HIV) *vif* compromise virus replication in human T cell lines that stably express the restriction factors APOBEC3F (A3F) or APOBEC3G (A3G). We previously reported that partial resistance to A3G could develop in these Vif-deficient viruses through a nucleotide A200-to-T/C transversion and a *vpr* null mutation, but these isolates were still susceptible to restriction by A3F. Here, long-term selection experiments were done to determine how these A3G-selected isolates might evolve to spread in the presence of A3F. We found that A3F, like A3G, is capable of potent, long-term restriction that eventually selects for heritable resistance. In 7/7 instances, the selected isolates had restored Vif function to cope with A3F activity. In two isolates, Vif Q26-Q27 and Y26-Q27, the resistance phenotype recapitulated in molecular clones, but when the selected *vif* alleles were analyzed in the context of an otherwise wildtype viral background a different outcome emerged. Although HIV clones with Vif Q26-Q27 or Y26-Q27 were fully capable of overcoming A3F, they were now susceptible to restriction by A3G. Concordant with prior studies, a lysine at position 26 proved essential for A3G neutralization. These data combine to indicate that A3F and A3G exert at least partly distinct selective pressures and that Vif function may be essential for the virus to replicate in the presence of A3F.

Introduction

Human APOBEC3F (A3F) is a DNA cytosine deaminase that, like its prototypical relative APOBEC3G (A3G), restricts the infectivity of human immunodeficiency virus type 1 (HIV) in the absence of the viral accessory protein virion infectivity factor (Vif) (Bishop *et al.*, 2004; Liddament *et al.*, 2004; Sheehy *et al.*, 2002; Wiegand *et al.*, 2004; Zheng *et al.*, 2004). Vif is thought to permit productive HIV infection by mediating the proteasomal degradation of A3G and A3F through recruitment of a ubiquitin ligase complex consisting of Elongins B and C, Cullin 5, and Rbx2 (Conticello *et al.*, 2003; Liddament *et al.*, 2004; Marin *et al.*, 2003; Mehle *et al.*, 2004a; Mehle *et al.*, 2004b; Sheehy *et al.*, 2003; Stopak *et al.*, 2003; Wiegand *et al.*, 2004; Yu *et al.*, 2003; Yu *et al.*, 2004c; Zheng *et al.*, 2004). In addition, several groups have suggested other mechanisms by which Vif may counteract APOBEC3 proteins, including directly inhibiting packaging, decreasing translation, and inhibiting deaminase activity (Kao *et al.*, 2003; Mariani *et al.*, 2003; Mercenne *et al.*, 2009; Santa-Marta *et al.*, 2005; Stopak *et al.*, 2003).

In contrast with HIV, direct and indirect observations indicate that many retroviruses do not need Vif to evade restriction by cellular APOBEC3 proteins. For instance, the lentivirus equine infectious anemia virus (EIAV) lacks a *vif* gene despite the existence of an extensive repertoire of equine APOBEC3 proteins (Bogerd *et al.*, 2008; Zielonka *et al.*, 2009), and among other types of Vif-deficient retroviruses, a number of alternative APOBEC3-resistance mechanisms have evolved. Foamy viruses use an auxiliary protein called Bet to neutralize APOBEC3 proteins (Löchelt *et al.*, 2005; Russell *et al.*, 2005). Human T cell leukemia virus type 1 (HTLV-1) avoids APOBEC3

encapsidation through its unique nucleocapsid protein (Derse *et al.*, 2007). Analogously, murine leukemia virus (MLV) excludes certain APOBEC3 proteins from virions and cleaves encapsidated murine APOBEC3 (Abudu *et al.*, 2006; Doehle *et al.*, 2005b), while Mason-Pfizer Monkey Virus (MPMV) also avoids the encapsidation of its cognate rhesus macaque A3G (Doehle *et al.*, 2006).

Overall, it appears that most retroviruses must possess a mechanism to evade the APOBEC3 proteins of their hosts, and this requirement does not always involve Vif. Even in the case of HIV, the most influential study to date to analyze the activity of primary isolates of Vif against A3F and A3G demonstrated that 10/40 genetically intact *vif* alleles tested were defective for the neutralization of one or both restriction factors (Simon *et al.*, 2005). In addition to intact but defective *vif* alleles, diverse studies suggest the occurrence of *vif* alleles with gross genetic lesions to be quite common, ranging from 3.4-31% of clones analyzed [*e.g.* 3.4% (Komoto *et al.*, 2005), 5.8% (Zhang *et al.*, 1997), 6.3% (Simon *et al.*, 2005), 10.2% (Yedavalli *et al.*, 1998), 17.4% (Sova *et al.*, 1995), 21.1% (Yedavalli and Ahmad, 2001) or 31.1% (Tominaga *et al.*, 1996)] and readily detectable in most studies in a sizeable fraction of patients analyzed [*e.g.* 0/50 (Hassaine *et al.*, 2000), 6/55 (Wieland *et al.*, 1994), 2/18 (Wieland *et al.*, 1997), 3/14 (Rangel *et al.*, 2009), 2/9 (Zhang *et al.*, 1997), 2/7 (Simon *et al.*, 2005), 4/10 (Sova *et al.*, 1995), 7/10 (Yedavalli *et al.*, 1998), 3/4 (Yedavalli and Ahmad, 2001) or 5/6 (Tominaga *et al.*, 1996)]. Our own analysis of 2,522 subtype B sequences from the Los Alamos National Laboratory HIV Sequence Database indicated approximately 4% of *vif* sequences contain one or more premature termination codons (data not shown; http://www.hiv.lanl.gov/). One may therefore conservatively estimate that, concordant with the overall proportion of

defective *vif* alleles estimated by Simon and colleagues, at least 20% of *vif* alleles are likely inactive against one or more APOBEC3 proteins (Simon *et al.*, 2005). Beyond that, some unknown percentage will likely display suboptimal neutralizing activity relative to "wildtype" as also observed previously [*e.g.* (Marin *et al.*, 2008; Simon *et al.*, 2005)]. Combining these findings with the fact that HIV has had only decades in which to adapt to the human APOBEC3 repertoire while its Vif-encoding simian immunodeficiency virus (SIV) ancestors may have been coexisting with primate APOBEC3 repertoires for millions of years [*e.g.* (Gifford *et al.*, 2008; Worobey *et al.*, 2008)], it is important to consider the possibility that Vif function may not be the sole mechanism of resistance to APOBEC3 proteins available to HIV.

To address this, we have previously carried out long-term *in vitro* selection experiments in which we passaged HIV containing tandem stop codons at position 26 and 27 of *vif* in the presence of A3G. These studies yielded viral variants that could spread effectively in the presence of A3G despite retention of these tandem stop codons; rather, all isolates had acquired a pyrimidine at nucleotide 200 and a stop codon in *vpr* (Haché *et al.*, 2008). While the effect of the *vpr* null mutation remains unknown, the substitution of A-to-T/C at nucleotide 200 functions to optimize viral translation, resulting in increased particle production and a relative decrease in A3G encapsidation to levels tolerable to the viral population (Haché *et al.*, 2009). However, we also observed that these T/C200 *vif-vpr-* variants were still susceptible to restriction by A3F and by the APOBEC3 repertoire in nonpermissive CEM cells (Haché *et al.*, 2008).

We therefore set out to select variants of these viruses that might additionally resist A3F. Given the literature on the Vif-independent evasion of APOBEC3 proteins by

the other viruses noted above as well what is known of the mechanisms of human APOBEC3-mediated restriction, we hypothesized that any of several possibilities might manifest: i) changes in nucleocapsid or viral nucleic acid structures that inhibit A3F encapsidation [*e.g.* references above and many addressing the mechanism of A3G encapsidation as discussed in (Chiu and Greene, 2008)]; ii) mutations in integrase or viral nucleic acid motifs that impede the effects of APOBEC3 proteins on integration [*e.g.* (Luo *et al.*, 2007; Mbisa *et al.*, 2007; Mbisa *et al.*, 2010)]; iii) mutations in viral nucleic acid structures or nucleocapsid that impede APOBEC3 effects on reverse transcription [*e.g.* (Bishop *et al.*, 2008; Guo *et al.*, 2006; Iwatani *et al.*, 2007; Li *et al.*, 2007; Yang *et al.*, 2007b)]; iv) nonspecific fitness-enhancing mutations [*e.g.* T/C-200 in (Haché *et al.*, 2008)]; or v) the restoration of functional *vif* alleles.

Results

A3F selects for the restoration of Vif function – To determine how *vif*-deficient HIV$_{IIIB}$ may evolve to overcome restriction by A3F, we used four starting virus templates to initiate a total of 288 selection cultures across four independent experiments. In the first two experiments, we used viruses derived from molecular clones of our previous A3G-resistant isolates to infect A3F-expressing CEM-SS cells [*i.e.* A3G-R2 is HIV$_{IIIB}$ T200 *vif*-X26X27 *vpr*-W54X and A3G-R3 is T200 *vif*-X26X27 *vpr*-W18X (Haché *et al.*, 2008); **Tables 2-1** and **2-2**]. Within approximately 8-12 weeks of long-term continuous culture, a total of 7 resistant isolates out of 96 infected were obtained; representative isolates from Experiments 1 and 2 are described below and in **Table 2-2**. All attempts to select resistance from the parental A200 *vif*-X26X27 *vpr*+ or a C200 derivative of this parent in these and in two subsequent experiments were unsuccessful (data not shown). Immunoblotting showed that the A3F levels in stably transfected clones of CEM-SS and the naturally nonpermissive CEM and H9 cell lines are similar [**Figure 2-1A** and (Haché *et al.*, 2008)]. A recent report further indicates that A3F expression levels in CEM are similar to those in primary CD4+ T cells, suggesting that the selective pressure applied in these studies is comparable to that exerted by A3F in primary cells (Refsland *et al.*, 2010).

A3F-resistant isolates were passaged sequentially two additional times in A3F-expressing CEM-SS cells. Replication curves from the third passage indicated that the resistant viruses spread with improved titer and kinetics in A3F-expressing cell lines with peaks generally occurring between 10 and 20 days post-infection as shown for 2 representative isolates (**Figure 2-1B**). As anticipated, the putative A3F-resistant isolates

maintained the ability to replicate in A3G-expressing or vector control CEM-SS clones (**Figures 2-1C** and **2-1D**). However, these viruses did not acquire the capacity to replicate in CEM (**Figure 2-1E**, below, and **Discussion**). In some instances, the growth curves of the parental virus (A200 Vif-K26-H27) are relatively weak due to high cytotoxicity, but this does not affect our overall interpretations or conclusions (*e.g.*, **Figure 2-1D**).

To genotype the A3F-selected viruses, we amplified and cloned a region of the proviral genome encompassing *vif* and *vpr*, nucleotides 3,095-5,485 [GenBank accession EU541617 (Haché *et al.*, 2008)]. We obtained a minimum of 10 sequences per isolate, and analysis revealed that the A3F-resistant isolates had changed nonsense *vif* codons to missense, while the *vpr* null mutations remained intact (**Table 2-2** and data not shown). A composite analysis of the sequences derived from three A3F-resistant isolates showed a high hypermutation load averaging 2.4 G-to-A mutations per kilobase (**Figure 2-2A-C**). These mutations generally occurred with notable bias toward a 5'-<u>G</u>A-3' dinucleotide context, which is characteristic of the deaminase activity of A3F attacking cDNA strand 5'-T<u>C</u>-3' motifs. However, substantial 5'-<u>G</u>G-3' context hypermutations consistent with the action of A3G were also evident, particularly in A3F-resistant isolate 7 (**Figure 2-2C**). These G-to-A mutations were likely the result of the low levels of endogenous A3G in CEM-SS cells [*e.g.* **Figure 2-1A**, see below and published quantifications (Koning *et al.*, 2009; Refsland *et al.*, 2010)]. The overall hypermutation levels in these isolates was also likely influenced by the existence of an initially mixed resistant population due to the sequential accumulation of the missense *vif* mutations, with codon 26 mutating first and codon 27 second. This order of events is suggested by the fact that the same X-to-Q

missense mutation at codon 26 was detected in isolate 7 alongside two different codon 27 missense mutations, X-to-Q or X-to-W. From hereon, analyses will focus exclusively on two resistance-associated *vif* alleles encoding missense mutations at codons 26 and 27, *vif*-QQ and *vif*-YQ (**Tables 2-1 & 2-2**).

Resistance to A3F occurs through Vif – Based on our proviral DNA sequence analyses and the fact that numerous prior studies have shown that Vif can counteract A3F, we hypothesized that the resistance to A3F observed in our isolates would be conferred by the aforementioned missense mutations at *vif* codons 26 and 27. To test this hypothesis and eliminate the possibility that mutations elsewhere in the viral genome might contribute to A3F resistance, we incorporated the *vif*-QQ and *vif*-YQ alleles cleanly into their parent A3G-resistant molecular clone backgrounds and infected A3F-expressing CEM-SS cells with viruses produced from these molecular clones (**Table 2-1**). These viruses – HIV$_{IIIB}$ T200 *vif*-QQ (or *vif*-YQ) *vpr-* – displayed robust infectivity and kinetics on cells expressing A3F, A3G, or a vector control, as expected for a virus engineered to contain both A3G and A3F resistance mutations (**Figures 2-3A-C**). In addition, like the original A3F-resistant isolates, these molecular clones still failed to replicate in CEM cells despite their enhanced replicative capacity in the presence of A3F (**Figure 2-3D**, below, and **Discussion**).

vpr deficiency does not explain the lack of resistant virus replication in CEM – Given the ability of the selected viruses to grow in the presence of A3F or A3G, it was notable that they were still unable to spread in CEM cells that express these and three additional

APOBEC3 proteins (Refsland *et al.*, 2010). We therefore turned to viral genotypes in search of a possible explanation. The nucleotide at position 200 was not likely to be responsible because the vast majority of HIV isolates already have a T or C at that position [*e.g.*, (Haché *et al.*, 2008) per the Los Alamos National Laboratory HIV Sequence Database]. The A200 in the parent HIV$_{IIIB}$ is presumably disfavored because it causes suboptimal translation, lower levels of viral particle production, and a relatively greater degree of A3G-dependent restriction (Haché *et al.*, 2009; Haché *et al.*, 2008). However, the function served by *vpr* inactivation in the parent A3G-resistant viruses is still unknown.

To address the possibility that an undescribed Vpr function might be required for these viruses to grow in the presence of multiple APOBEC3 proteins, we incorporated the selected *vif* alleles into HIV$_{IIIB}$ C200 *vpr*(+) molecular clones (**Table 2-1**). All viruses were replication-competent in vector control CEM-SS and SupT11 cell lines (data not shown). Like the Vpr-deficient *vif*-QQ and *vif*-YQ viruses in **Figures 2-1E** and **2-3D**, however, Vpr-proficient viruses with the *vif*-QQ or *vif*-YQ alleles remained restricted by nonpermissive CEM and H9 cells (**Figure 2-4A-B**, data not shown, and see below for additional data with these molecular clones). In contrast, the wildtype *vif*-KH virus spread readily in these fully nonpermissive lines in both *vpr+* and *vpr-* backgrounds (**Figure 2-4A-B** and data not shown). We therefore concluded that the basis for the observed replication defect in naturally nonpermissive cell lines lay in the identity of the amino acids selected by A3F – *vif*-QQ or *vif*-YQ versus the wildtype *vif*-KH.

Variants of Vif selected by A3F are unable to effectively degrade A3G – To mechanistically characterize Vif-QQ and Vif-YQ, we assessed the ability of these Vif variants to neutralize A3F or A3G in single-cycle infectivity experiments at several levels of Vif expression. As predicted by the spreading infection phenotypes, the Vif variants selected by A3F were as efficient at enhancing viral infectivity and lowering intracellular A3F levels as the wildtype Vif-KH (**Figure 2-5A**). In contrast, the Vif variants selected by A3F did not restore infectivity in the presence of A3G, and they were markedly deficient in their ability to lower intracellular A3G levels (**Figure 2-5B**). This clear A3G susceptibility helped to explain why viruses with these *vif* alleles fail to spread in CEM or H9 (*e.g.* **Figure 2-4**) and why our original A3F-resistant isolates had significant 5'-G̲G̲-3' context hypermutations – the original CEM-SS A3F lines have permissive but detectable levels of endogenous A3G [**Figure 2-1A**, **Figure 2-2** and (Refsland *et al.*, 2010)]. However, it remains possible that other APOBEC3 proteins expressed in CEM and H9 beyond A3F and A3G may account for at least part of the continued restriction of A3F-selected viruses in these cells [**Discussion** and (Refsland *et al.*, 2010)].

The inability of the vif alleles selected by A3F to neutralize A3G maps to the identity of amino acid 26 – To test whether viruses with the *vif* alleles selected by A3F are deficient in their ability to neutralize A3G in spreading infections and, if so, to additionally determine the nature of the mutation responsible, we did a series of spreading infection experiments using a panel of full-length molecular clones with amino acid substitution mutations at *vif* codons 26 or 27 (*i.e.* wildtype HIV_{IIIB} C200 viruses with the exception of *vif*; **Table 2-1**). Specifically, we compared viruses with K, Q, or Y at position 26 and H

or Q at position 27, and we used A3F or A3G expressing SupT11 clones because we recently found that this line is nearly devoid of endogenous APOBEC3 expression [(Refsland *et al.*, 2010) and **Figure 2-6A**].

We found that all combinations of Vif residues at positions 26 and 27 were able to spread in SupT11 cells expressing A3F and in SupT11 vector control cells (**Figure 2-6B** and data not shown). In contrast, only viruses encoding a lysine at position 26 of Vif were able to spread effectively in cells expressing A3G or in nonpermissive CEM or H9 cell lines (**Figures 2-6C-D** and data not shown). These data corroborate two recent Vif site-directed mutation screens, which independently found that K26 is important for the neutralization of A3G (Chen *et al.*, 2009; Dang *et al.*, 2009). They are also consistent with the high conservation of K26 in sequences in the Los Alamos National Laboratory HIV sequence database, where 99.4% (2,507/2,522) of subtype B sequences encode lysine at position 26 [data not shown; http://www.hiv.lanl.gov/].

Given the correlation between the capacity of a particular Vif to degrade APOBEC3 proteins and the ability of viruses carrying those *vif* alleles to spread in the presence of APOBEC3 proteins, we further hypothesized that degradation is the predominant mechanism of Vif action at play in our spreading infection system. To test this hypothesis, we incorporated Vif BC Box mutations at positions 144-146 into both the wildtype and A3F-selected molecular clones. SLQ-to-AAA substitutions in this highly conserved region have been previously shown to interrupt interaction between Vif and Elongin C, thus attenuating the Vif-mediated degradation of APOBEC3 proteins [*e.g.* (Yu *et al.*, 2003)]. We observed that viruses carrying these mutations are unable to spread in SupT11 cells stably transfected with A3F or A3G as well as naturally nonpermissive

CEM or H9 cells (data not shown). These data indicate that the interaction between Vif and Elongin C is required for A3F/G neutralization and, further, based on many prior studies, that degradation is an integral part of the mechanism.

Additional long-term selection experiments in nonpermissive CEM cells select for a positive charge at Vif position 26 – To determine how the *vif* alleles selected by A3F might evolve to become resistant to the full nonpermissive APOBEC3 repertoire in CEM, we passaged *vif*-QQ viruses in these cells until resistant isolates arose. From 48 parallel cultures, 5 isolates emerged after 8-12 weeks of incubation. Second passage proviral DNA sequencing revealed that 3 isolates carried Q26R mutations while two more carried Q26K mutations; representative viruses from these selections are shown in Experiment 5 of **Table 2-2**. Serial passage in CEM indicated that these isolates replicated with greater efficiency than the parent *vif*-QQ viruses, although Q26R viruses still tended to display attenuated replication kinetics (**Figure 2-7**).

Discussion

The experiments described here are the first reported selections for HIV resistance to the restriction factor A3F. Our data show that A3F, like A3G, is capable of potent long-term restriction, extending to two the number of APOBEC3 proteins shown capable of selecting heritable resistance (Haché *et al.*, 2008). The strength of A3F-mediated restriction in our spreading infection experiments combined with its ability to select resistance contrasts somewhat with the consistently weaker effects of A3F in single-cycle assays compared with A3G [*e.g.* (Bishop *et al.*, 2006; Bishop *et al.*, 2004; Holmes *et al.*, 2007; Liddament *et al.*, 2004; Simon *et al.*, 2005)]. The reason(s) for this disparity are not obvious, but the data presented here suggest that A3F may be an important part of the selective pressure that maintains Vif function *in vivo*.

We presume that the isolates emerging from long-term selection experiments will have taken the easiest path to resistance. Given the current literature on APOBEC3 proteins, we might have anticipated the selection of any number of Vif-independent resistance mechanisms to A3F as discussed in the **Introduction**, but we have only ever selected for the restoration of Vif function. If Vif-independent resistance to A3F is possible, then, the genetic barriers to evolving such mechanisms are apparently greater than the 1-2 mutational events required to restore Vif function in the alleles described here. Fitness costs to viruses acquiring these putative alternative resistance mutations may also influence the overall barrier to their development.

Aside from the central observation that A3F selects for the restoration of Vif function, we identified K26 as an important residue for the Vif-mediated degradation of A3G but not A3F. This result is satisfying in light of recent site-directed mutagenesis

85

studies also showing that K26 is indeed required for A3G neutralization (Chen *et al.*, 2009; Dang *et al.*, 2009), and it likely explains the notable presence of G-to-A mutations consistent with A3G action in our selected viral isolates (**Figures 2-1A** and **2-2**). Molecularly, our characterization of the *vif* alleles selected by A3F is largely in agreement with these two reports in finding that viruses without the wildtype K26 lack the ability to efficiently degrade A3G, although they still continue to bind it (**Figure 2-5B** and data not shown). Thus, our data also support the notion that K26 is part of the Vif interaction surface that contacts A3G.

In the process of characterizing the amino acid identity requirements at residue 26, we also demonstrate for the first time that the evolution of *vif* in response to one APOBEC3 protein (A3F) may not result in a Vif variant with activity against another APOBEC3 protein (A3G; **Figures 2-5** and **2-6**). Thus, the selective pressures exerted on *vif* by each APOBEC3 protein are at least partially distinct. This conclusion is consistent with a growing body of mutagenesis data implying that the structural determinants of A3F and A3G binding sites within Vif can be both defined and in some instances separated genetically [reviewed in (Albin, 2010; Smith *et al.*, 2009)].

While we have not yet succeeded in selecting Vif function as a mechanism of A3G resistance, our observations should not be extrapolated to quantify the relative abilities of A3F and A3G to select for Vif function. For instance, the *vif* genotype of the parental virus is highly likely to influence the outcome. Furthermore, we do not wish to imply that Vif function is unimportant for overcoming restriction by A3G. Even *in vitro*, the A3G-selected mutations that we have previously reported are not as robust as wildtype Vif function. They fail to complement the A3F-selected Vif alleles when

replicated in naturally nonpermissive CEM and H9 cells as well as CEM-SS cells stably transfected with both A3F and A3G (**Figures 2-1E, 2-3D** and data not shown), and we have also observed that they may become overwhelmed when replicating on SupT11 cells expressing higher levels of A3G but no other APOBEC3 proteins (data not shown). This is likely due to the nature of our previously characterized A3G-resistance mutations, which merely tolerate the presence of A3G better than the parental virus (Haché *et al.*, 2008). That a particular resistance mutation should be saturable by higher APOBEC3 expression levels is perhaps not all that surprising, because even Vif-proficient viruses can be overwhelmed and restricted under conditions of APOBEC3 overexpression [*e.g.* (Peng *et al.*, 2006; Sadler *et al.*, 2010)]. Furthermore, we cannot rule out the possibility that other dominant acting APOBEC3 restriction factors in CEM and H9 beyond A3F and A3G may be at least in part responsible for the more potent restriction observed in these cell lines. Results of future long-term selection experiments, especially if the viral genotype and/or the cellular APOBEC3 levels differ from those described here, are likely to yield other informative outcomes.

Materials and Methods

Plasmids – All novel constructs described were confirmed by DNA sequencing. A3F and A3G coding sequences are identical to those found in GenBank, NM_145298 and NM_021822, respectively. The pcDNA3.1-derived constructs used for the stable expression of untagged A3F and A3G have been described (Haché *et al.*, 2008). pcDNA3.1-V5 and pcDNA3.1-A3F-V5 expression constructs containing a V5 epitope tag were made by digesting pcDNA3.1-3xHA or pcDNA3.1-A3F-3xHA (Stenglein and Harris, 2006) with XhoI/XbaI and ligating these with synthetic complementary oligonucleotides 5'-TC GAG GGA GTC GAG GGC GGC GGT AAG CCT ATC CCT AAC CCT CTC CTC GGT CTC GAT TCT ACG TAG T-3' and 5'-CT AGA CTA CGT AGA ATC GAG ACC GAG GAG AGG GTT AGG GAT AGG CTT ACC GCC GCC CTC GAC TCC C-3', which replaces the 3xHA coding sequence with one for the V5 epitope. pcDNA3.1-A3G-V5 was derived by amplifying A3G from pcDNA3.1-A3G using primers 5'-NNN GAA TTC GAG CTC GGT ACC ACC ATG AAG CCT CAC TTC AG AAA C-3' and 5'-NNN GTC GAC TCC GTT TTC CTG ATT CTG GAG AAT-3', digesting with KpnI/SalI and ligating the purified product into pcDNA3.1-V5 digested with KpnI/XhoI.

We codon-optimized HIV$_{IIIB}$ Vif and added nucleotides for a C-terminal HA epitope (GenScript USA). The translated Vif open reading frame is identical to GenBank EU541617 (Haché *et al.*, 2008). Vif$_{IIIB}$-HA was excised from parental vector pU57 using BamHI/SalI and subcloned into a similarly cut pVR1012-derived plasmid provided by Dr. Xiao-Fang Yu (Johns Hopkins University, Baltimore). This BamHI/SalI-digested expression construct was religated to itself to create a vector control. pVR1012-Vif$_{IIIB}$-

HA was modified by site-directed mutagenesis with Pfu polymerase (Stratagene) to create variants encoding QQ and YQ at positions 26 and 27 using primers 5'-T TGG AAG CGC CTC GTG CAG CAG CAT ATG TAC ATC TCC C-3' and 5'-G GGA GAT GTA CAT ATG CTG CTG CAC GAG GCG CTT CCA A-3' or 5'-ACT TGG AAG CGC CTC GTG TAT CAG CAT ATG TAC ATC TCC CGC-3' and 5'-GCG GGA GAT GTA CAT ATG CTG ATA CAC GAG GCG CTT CCA AGT-3'.

Proviral constructs for wildtype HIV_{IIIB} [GenBank EU541617 (Haché *et al.*, 2008)] and a *vif*-deficient derivative with tandem stop codons at positions 26 and 27 were obtained from Dr. Michael Malim (King's College London, England). Derivative molecular clones with A3G-resistant virus 2 and A3G-resistant virus 3 (A3G-R2 and A3G-R3) mutations, as well as a *vif*-deficient HIV_{IIIB} containing a cytosine at nucleotide 200, have been described [(Haché *et al.*, 2008) and **Table 2-1**]. Substitution of a given *vif* allele in these backgrounds was done by site-directed mutagenesis on a pCR4-Blunt (Invitrogen) plasmid containing the PCR amplified and cloned HIV_{IIIB} *vif/vpr* region extending from nucleotides 3,095 to 5,485 [GenBank EU541617 (Haché *et al.*, 2008), where +1 is the transcriptional start site]. Of note, the A3G-R3 used here is a variant of the published sequence containing an additional nucleotide deletion at position 5,285 in the *vpr* open reading frame downstream of the reported *vpr* W18X mutation (Haché *et al.*, 2008). Mutated fragments were reintroduced into full-length molecular clones by subcloning the SwaI/SalI fragment from the shuttle vector into a similarly cut parental HIV_{IIIB} proviral plasmid. A comprehensive listing of the molecular clone genotypes used in these studies is provided in **Table 2-1**.

Stable cell lines – CEM-SS cells stably expressing A3F, A3G or a vector control have been described (Haché *et al.*, 2008). SupT1 was obtained from the AIDS Research and Reference Reagent Program (#100). SupT11 is a single-cell subclone of SupT1 isolated by limiting dilution. APOBEC3-expressing derivatives were made by electroporating with 20 µg linearized plasmid (250 volts, 950 µF; BioRad), selecting with 1.0 mg/ml G418 (Mediatech), and screening clones for APOBEC3 expression by immunoblotting using anti-A3F or anti-A3G antibodies (#11474 or #10201 from Drs. Michael Malim or Jaisri Lingappa, respectively, obtained through the AIDS Research and Reference Reagent Program). T cell lines were maintained in RPMI supplemented with 10% fetal bovine serum, penicillin/streptomycin, β-mercaptoethanol and, in some cases, 0.5 mg/mL G418.

Virus stocks – Infectious virus was produced by transfecting 5 or 10 µg of proviral molecular clone plasmids into 293T cells at approximately 70% confluence in 10 cm dishes using Trans-IT (Mirus Bio) or Fugene 6 (Roche) transfection reagents. 2-3 days later, virus-containing supernatants were harvested and filtered through 0.45 µm filters. 293T cells were maintained in DMEM supplemented with 10% fetal bovine serum and penicillin/streptomycin.

Virus titration – Viruses derived from molecular clones were titered by infecting 50,000 CEM-GFP reporter cells (Gervaix *et al.*, 1997; Haché *et al.*, 2008). Approximately three days later, cells were fixed in 4% paraformaldehyde, and GFP-positive infected cells were quantified using flow cytometry. A Beckman-Coulter Quanta MPL or a Becton-

Dickinson FacsCalibur instrument was used for data collection, and FlowJo (Tree Star) or CellQuest (Becton-Dickinson) software was used for data analysis. Live cells were gated by forward versus side scatter (or by electronic volume versus side scatter when using the Quanta MPL), and the proportion of GFP-positive cells in this population was quantified. These percentages were then graphed against the volume of virus stock used, and linear regression was employed to determine the percentage of cells infected at a given volume of a particular virus stock. Calculations of multiplicity of infection (MOI) on CEM-GFP cells were derived from these percentages and used as a standard by which to normalize the quantity of infectious virus used to initiate infections. Quantification of resistant virus isolates differed from this protocol due to the limited amount of passaged supernatants available. In this case, titers were calculated from single infections of 25,000 CEM-GFP cells with 150 μL of virus saved on a given day in a total volume of 250 μL.

HIV spreading infection experiments – Spreading infections were initiated by infecting 50,000 cells of a given cell line in a total volume of 1 mL in one well of a 24-well plate at a CEM-GFP MOI of 0.01-0.05; MOI in a given experiment is indicated in the associated figure legend. Spreading infections were monitored by periodically using 150 μL of cell-free supernatant to infect fresh target CEM-GFP cells at a concentration of 25,000 cells per 100 μL per well of a 96-well plate. At approximately 2-4 days post-infection, these cells were fixed in 4% paraformaldehyde and analyzed by flow cytometry as above. The original infected cultures were split and their media replenished as needed at each time point taken to prevent the overgrowth of infected cells. The accumulation of multiple cycles of infection in a spreading infection leads to some variability in the peaks of

91

replication, and clonal variation among the cell lines used in the experiments described introduces an additional variable (See *Stable Cell Lines* above for cell line derivation procedures). Nevertheless, general trends of growth/no growth for each virus on each cell line are apparent in the data shown here and in many additional experiments not shown. Additional observations of import are provided in the text or associated figure legend to supplement raw infectivity curves. Except where specifically indicated, the restriction of a given virus is indicated by increasingly flat replication curves. Viruses that spread effectively, in contrast, will generally yield a sharp infectivity peak followed by a decline as cells are killed by viral spread.

Resistant virus selection procedures – Resistant viruses were selected by initiating spreading infections in a given cell line (*e.g.* CEM-SS clones stably transfected with A3F) as indicated above at an MOI of 0.05 or 0.03. Cell-free supernatants were periodically used to infect target CEM-GFP cells as above, and these CEM-GFP cultures were visually monitored for the outgrowth of notably more infectious viruses as evidenced by an increased proportion of bright GFP-positive cells. Cultures were split and fed as needed at each monitoring point to prevent the overgrowth of infected cells on which viral spread remained restricted until the emergence of resistance. Upon outgrowth of a virus displaying enhanced infectivity, cell-free supernatants were saved at each monitoring point for subsequent confirmatory passage. Passage of resistant isolates was then initiated using virus produced in the presence of A3F, and these passages were conducted as indicated above for a typical spreading infection.

Proviral DNA sequencing – Genomic DNA was extracted from target CEM-GFP cells infected at the end of the third passage of resistant virus isolates using Qiagen DNeasy kits. This gDNA was then used as template for the amplification of integrated proviral sequences using PCR with high fidelity Phusion polymerase (NEB) and primers 5'-CAAGGCCAATGGACATATCA-3' and 5'-CAAACTTGGCAATGAAAGCA-3'. PCR products were cloned into pJET (Fermentas), TOPO pCR4-Blunt or TOPO pCR-Blunt II (Invitrogen). Cloned viral fragments were recovered in TOP10 or DH10B *Escherichia coli* and plasmid DNA prepared with Qiagen or Clontech miniprep kits. Plasmids were then sequenced using universal or virus-specific primers. Sequences were compiled and analyzed using Sequencher software (Gene Codes Corp.).

APOBEC3 immunoblotting experiments – T cell lines were grown to confluence, harvested, and lysed in approximately 50 µl of lysis buffer [25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 50 µM $ZnCl_2$, 10% glycerol and 1% Triton X-100 lysis buffer containing 50 µM MG132 (American Peptide) and complete protease inhibitor (Roche)] per 1 mL of cells. A 5x sample buffer consisting of 62.5 mM Tris pH 6.8, 20% glycerol, 2% sodium dodecyl sulfate (SDS), 5% β-mercaptoethanol and 0.05% bromophenol blue was added to each lysate to a final concentration of 2x, and the mixture was boiled for approximately 10 minutes prior to loading. Proteins were fractionated by 10% SDS-PAGE, transferred to a PVDF membrane, blocked in 4% milk dissolved in phosphate buffered saline (PBS) 0.01% Tween, incubated with rabbit anti-A3F (above), rabbit anti-A3G (above) or mouse anti-Tubulin (Covance) and probed with a secondary horseradish peroxidase (HRP)-conjugated goat anti-rabbit or goat anti-mouse antibodies (BioRad).

Blots were then developed using HyGlo chemiluminescent HRP antibody detection reagent (Denville Scientific) and exposed to film. Membranes were stripped using 62.5 mM Tris pH 6.8, 2% SDS and 100 mM β-mercaptoethanol at 50ºC and washed in PBS 0.01% Tween prior to sequential blocking and reprobing with each primary antibody after the first.

Single-cycle infectivity assays – 250,000 293T cells/well were plated in a total volume of 2 ml/well in 6-well plates. Approximately 24h later, each well was transfected using Trans-IT transfection reagent (Mirus Bio) with 200 ng of a vector control or an APOBEC3-V5 expression construct, a pVR1012 vector control or 50, 100 or 200 ng of a given Vif-HA expression construct supplemented as necessary to 200 ng total with the pVR1012 vector and 1.6 µg of the full-length replication-competent HIV$_{IIIB}$ C200 XX provirus (**Table 2-1**). Two days post-transfection, virus-containing supernatants were filtered through 0.45 µm filters and used to infect CEM-GFP cells in 96-well plates as above. Three days post-infection, these CEM-GFP cells were fixed and analyzed by flow cytometry as above.

In addition to particle infectivity, producer cell APOBEC3 expression levels were determined. Cells were lysed in 25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM MgCl$_2$, 50 µM ZnCl$_2$, 10% glycerol and 1% Triton X-100 lysis buffer containing 50 µM MG132 (American Peptide) and complete protease inhibitor (Roche). Immunoblots were then prepared as described above, except primary mouse anti-V5 (Invitrogen), mouse anti-HA.11 (Covance) or mouse anti-Tubulin (Covance) antibodies were used for protein

detection in conjunction with secondary HRP-conjugated goat anti-mouse (BioRad) antibodies.

Table 2-1: HIV molecular clone genotypes used in this study.

Name HIV$_{IIIB}$	Nucleotide 200	*vif* Codons 26-27	*vpr*	Other Mutations	Reference
A200 XX	A	XX	+	n/a	Figure 1b-e; 3a-d
C200 XX	C	XX	+	n/a	Figure 4a-b; 5a-b; 6b-d; 7
A200 KH	A	KH	+	n/a	Figure 1b-e; 3a-d
C200 KH	C	KH	+	n/a	Figure 4a-b; 6b-d; 7
C200 KH SLQ>AAA	C	KH	+	*vif* 144-146 SLQ>AAA	Figure 6 data not shown
T200 XX *vpr-* (A3G-R2)	T	XX	W54X	*gag* (NC) R32K; *pol* (IN) silent	Figure 1b-e; 3a-d
T200 XX *vpr-* (A3G-R3)	T	XX	W18X	Del A5285 (*vpr*)	Figure 1b-e; 3a-d
T200 KH *vpr-* (A3G-R2)	T	KH	W54X	*gag* (NC) R32K; *pol* (IN) silent	Figure 4 data not shown
T200 KH *vpr-* (A3G-R3)	T	KH	W18X	Del A5285 (*vpr*)	Figure 4 data not shown
T200 QQ *vpr-* (A3F-R7)	T	QQ	W54X	*gag* (NC) R32K; *pol* (IN) silent	Figure 3a-d
T200 YQ *vpr-* (A3F-R5)	T	YQ	W18X	Del A5285 (*vpr*)	Figure 3a-d
C200 QQ	C	QQ	+	n/a	Figure 4a-b; 6b-d; 7
C200 QQ SLQ>AAA	C	QQ	+	*vif* 144-146 SLQ>AAA	Figure 6 data not shown
C200 YQ	C	YQ	+	n/a	Figure 4a-b; 6b-d
C200 YQ SLQ>AAA	C	YQ	+	*vif* 144-146 SLQ>AAA	Figure 6 data not shown
C200 KQ	C	KQ	+	n/a	Figure 6b-d; 7 data not shown
C200 QH	C	QH	+	n/a	Figure 6b-d
C200 YH	C	YH	+	n/a	Figure 6b-d
C200 RQ	C	RQ	+	n/a	Figure 7 data not shown

Table 2-2: Summary of experiments selecting A3F- or CEM-resistant HIV variants.

Expt	Cell Line	Parental HIV$_{IIIB}$ Genotype	Resistant Isolate Name	Selected *vif* 26-27 Genotype	Amino Acid 26-27 Changes	Frequency by Sequencing
		A200 KH *vpr+*	n/a	AAA CAC	KH	n/a
		A200 XX *vpr+*	n/a	TAA TAG	XX	n/a
1	CEM-SS F1	T200 XX *vpr* W54X (A3G-R2)	A3F-R3	CAA CAG	QQ	11/13
1	CEM-SS F1	T200 XX *vpr* W18X (A3G-R3)	A3F-R5	TAT CAG	YQ	12/13
2	CEM-SS F2	T200 XX *vpr* W54X (A3G-R2)	A3F-R7	CAA CAG / CAA TGG	QQ / QW	5/12 / 2/12
5	CEM	C200 QQ *vpr+*	CEM-R1	AAA CAG	KQ	5/5
5	CEM	C200 QQ *vpr+*	CEM-R2	CGA CAG	RQ	3/3
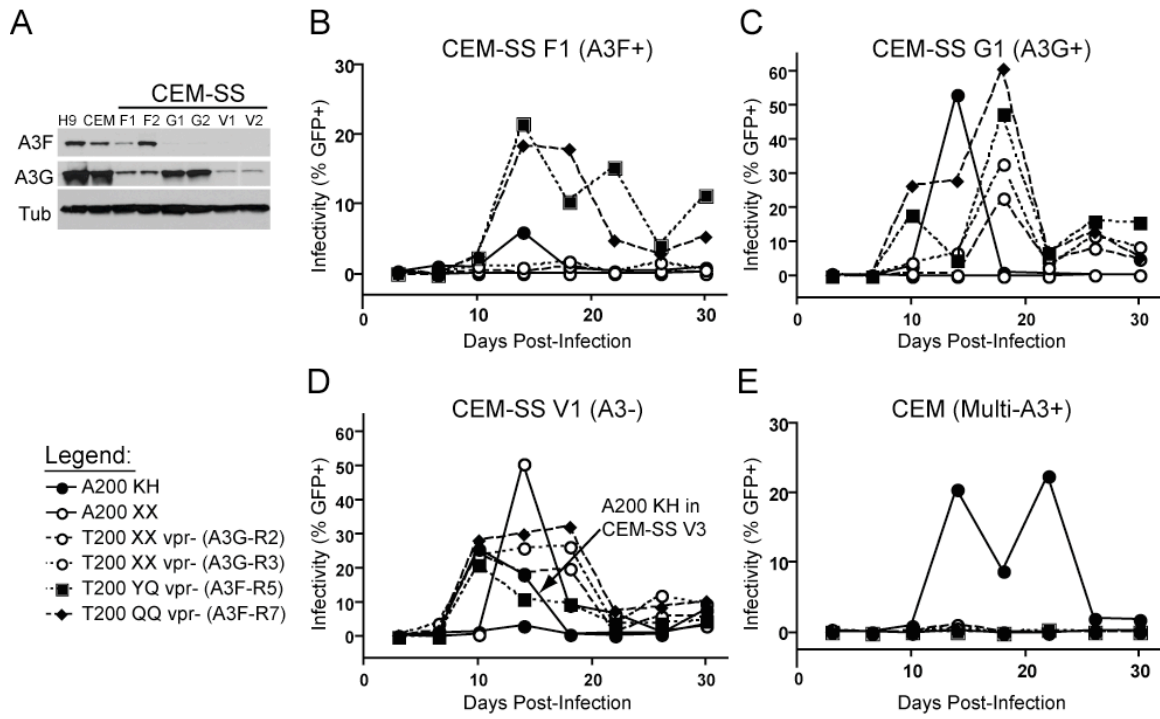
Figure 2-1: Restriction of Vif-deficient HIV by A3F selects resistant virus variants.

Figure 2-1: Restriction of Vif-deficient HIV by A3F selects resistant virus variants.

(**A**) A Western blot showing the expression of A3F and A3G in H9, CEM and the CEM-SS-derived cell lines used in these studies. F1 and F2 = A3F-expressing CEM-SS; G1 and G2 = A3G-expressing CEM-SS; V1 and V2 = CEM-SS transfected with a vector control. (**B-E**) Growth curves in the indicated cells for the following HIV isolates: Vif-proficient (A200 KH), Vif-deficient (A200 XX), A3G-resistant [(T200 XX *vpr-* (A3G-R2) and T200 XX *vpr-* (A3G-R3)] and representative A3F-resistant viruses derived from parent A3G-resistant viruses [T200 YQ *vpr-* (A3F-R5) and T200 QQ *vpr-* (A3F-R7) where *vif* genotypes are retrospectively indicated according to subsequent sequencing (**Table 2-2**)]. The starting MOI was approximately 0.02, and similar results were obtained using second CEM-SS A3F-expressing and vector control cell lines (data not shown). The low peaks observed for some A200 KH growth curves, particularly in panel **2-1D**, are due to the high cytotoxicity of this virus, which sometimes results in low apparent titers as infected cells are rapidly killed. In addition to the notably low A200 KH CEM-SS V1 curve, a growth curve for A200 KH in another CEM-SS vector line from this experiment, CEM-SS V3, is indicated by an arrow in **2-1D** to visually demonstrate that the wildtype virus spreads in the absence of restrictive levels of APOBEC3 proteins. The x-axis is offset from zero in all curves to permit better visualization of viruses showing little or no growth. Throughout, open symbols indicate viruses lacking Vif expression; closed symbols indicate full-length *vif* alleles. Similarly, dashed lines are used for *vpr-* viruses, while solid lines are used for *vpr+* viruses.

A

A3F-Resistant Isolate 3

	A	C	G	T
A		0	57	0
C	0		1	1
G	1	0		0
T	0	2	0	

G-to-A Dinucleotide Context (29.9 kb):
GA 70% GC 2% GG 26% GT 2%

B

A3F-Resistant Isolate 5

	A	C	G	T
A		0	20	0
C	0		0	0
G	2	0		0
T	1	13	0	

G-to-A Dinucleotide Context (29.9 kb):
GA 65% GC 5% GG 30% GT 0%

C

A3F-Resistant Isolate 7

	A	C	G	T
A		0	129	0
C	0		0	1
G	0	0		1
T	0	16	0	

G-to-A Dinucleotide Context (27.6 kb):
GA 37% GC 5% GG 58% GT 0%
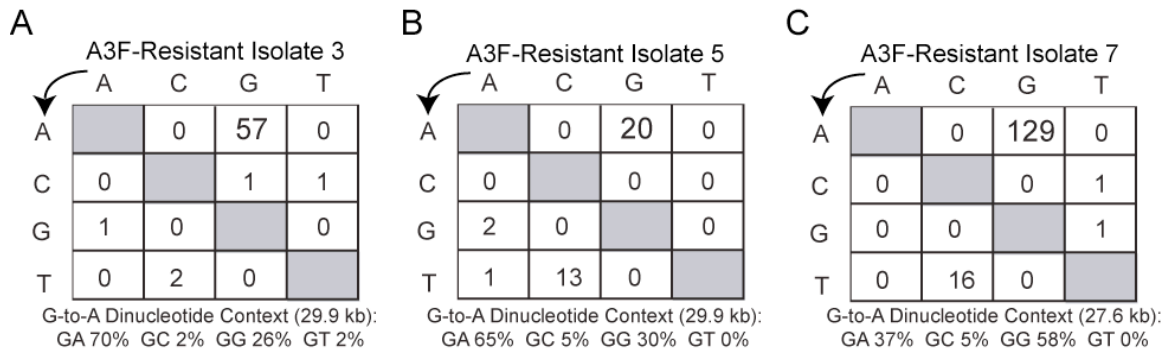
Figure 2-2: Hypermutation patterns in selected A3F-resistant isolates. The frequency of each base change is given for the clones described in Experiments 1 and 2 of **Table 2-2**, as is the predominance of the dinucleotide context in which G-to-A mutations occur. Similar to previous authors, we note a substantial C-to-T transition rate in the presence of A3F in addition to the expected G-to-A hypermutations [*e.g.* (Bishop *et al.*, 2004; Holmes *et al.*, 2007; Liddament *et al.*, 2004)]. The viruses selected also contain substantial 5'-GG-3' to 5'-AG-3' transitions, particularly in the case of A3F-resistant isolate 7, suggesting ongoing mutation by A3G in addition to A3F.
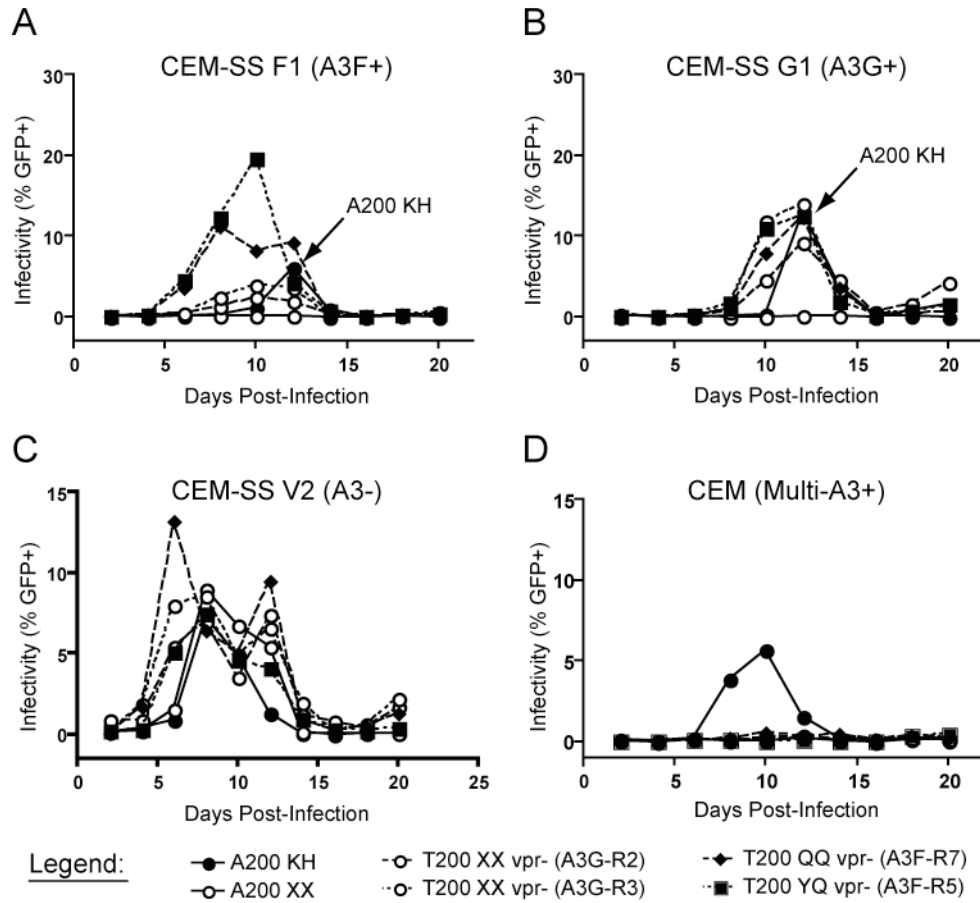
Figure 2-3: Restoration of the *vif* open reading frame accounts for phenotypic resistance to A3F-mediated restriction.

Figure 2-3: Restoration of the *vif* open reading frame accounts for phenotypic resistance to A3F-mediated restriction. Spreading infections at an MOI of 0.05 were initiated in CEM-SS cells stably transfected with A3F (**A**), A3G (**B**) or a vector control (**C**) as well as in nonpermissive CEM cells (**D**) using viruses derived from proviral molecular clones with the indicated genotypes. The mildly enhanced infectivity of A3G-resistant viruses relative to their parent A200 XX viruses in A3F-expressing CEM-SS cells is sometimes observed in experiments such as the one shown that starts from a higher MOI (compare the lower MOI in **Figure 2-1B** with the higher MOI in **Figure 2-3A**). In contrast, A3F-selected viruses consistently display robust peaks at any MOI in the presence of A3F. The peaks of A200 KH growth are indicated by arrows in **2-3A** and **2-3B** to differentiate them from the descending T200 YQ *vpr*- (A3F-R5) curve and the superimposed T200 YQ *vpr*- (A3F-R5) peak, respectively. Similar results were obtained using proviral molecular clones corresponding to other selected isolates as well as additional CEM-SS clones stably transfected with A3F, A3G or a vector control (data not shown).
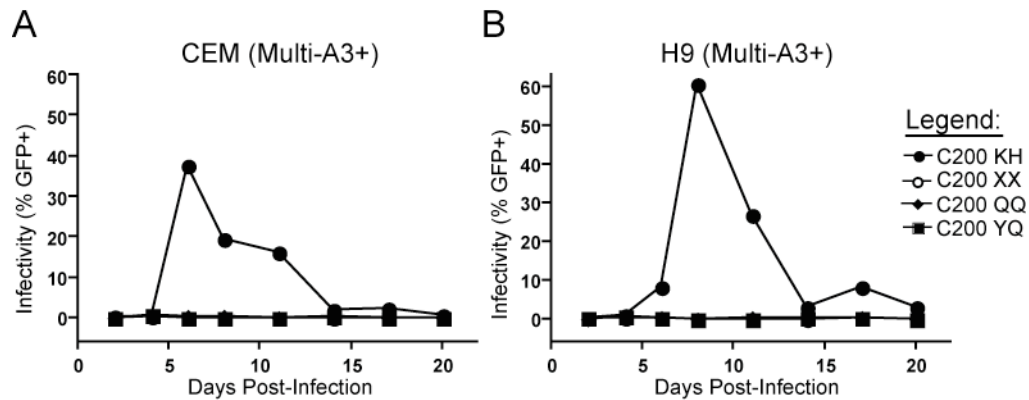
Figure 2-4: The identity of Vif amino acids 26 and/or 27 rather than Vpr status is critical for the ability to replicate on naturally nonpermissive cells. Spreading infection curves are shown for viruses with wildtype (KH), A3F-selected missense (QQ and YQ) and nonsense (XX) codons at positions 26 and 27 of *vif* in a Vpr-proficient context. Spreading infections were carried out from a starting MOI of 0.01 on CEM and H9 cells (**A** and **B**) as well as CEM-SS and SupT11 clones transfected with a vector control (data not shown). Different alleles of *vif* in Vpr-deficient contexts showed the same growth patterns on CEM and H9 as their Vpr-proficient counterparts (**Table 2-1** and data not shown).
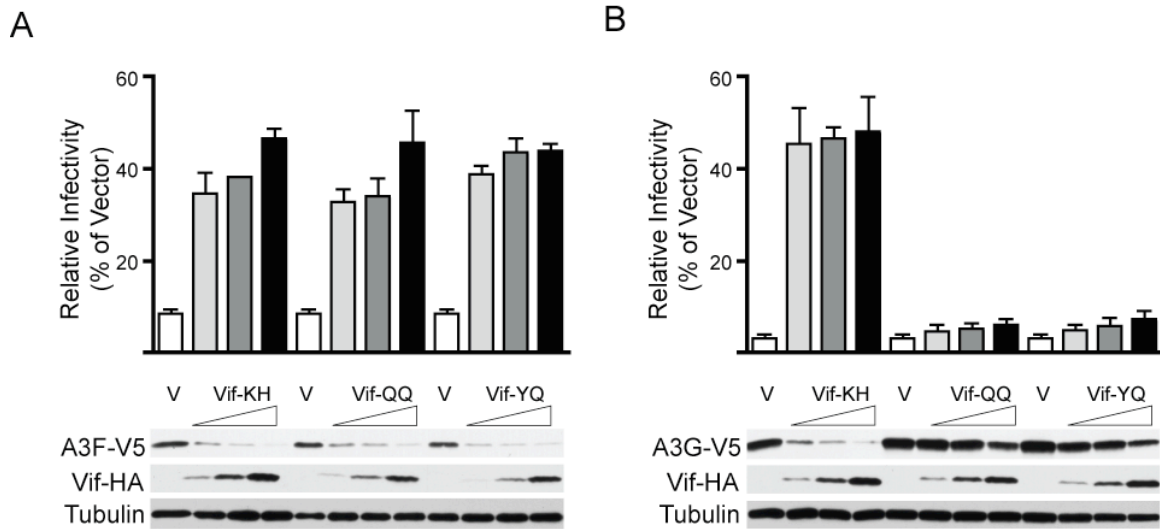
Figure 2-5: Functional Vifs selected by A3F are deficient in their ability to degrade A3G. A titration experiment analyzing the infectivity of particles produced by the cotransfection of constant amounts of A3F-V5 (**A**) or A3G-V5 (**B**) in the presence of increasing amounts of Vif-HA. The identities of amino acids 26 and 27 are indicated for each Vif-expressing construct. Infectivity data represent the mean and SEM of three independent experiments where infectivity is determined relative to that of particles produced under the same cotransfection conditions in each experiment with a vector control in place of the APOBEC3 expression construct (not shown). Immunoblots shown are taken from one of these three experiments. While Vif-QQ and Vif-YQ are notably deficient in their ability to neutralize A3G relative to the wildtype Vif-KH, a mild effect is seen at higher levels of Vif expression, which achieves statistical significance by a paired two-tailed t test for Vif-QQ but not Vif-YQ (**B** and data not shown). V = Vif vector and A3F-V5 or A3G-V5 expression constructs cotransfected; these conditions were tested once in each experiment but are loaded in the immunoblots and plotted in the histograms three times each for direct visual comparison with the addition of each Vif.
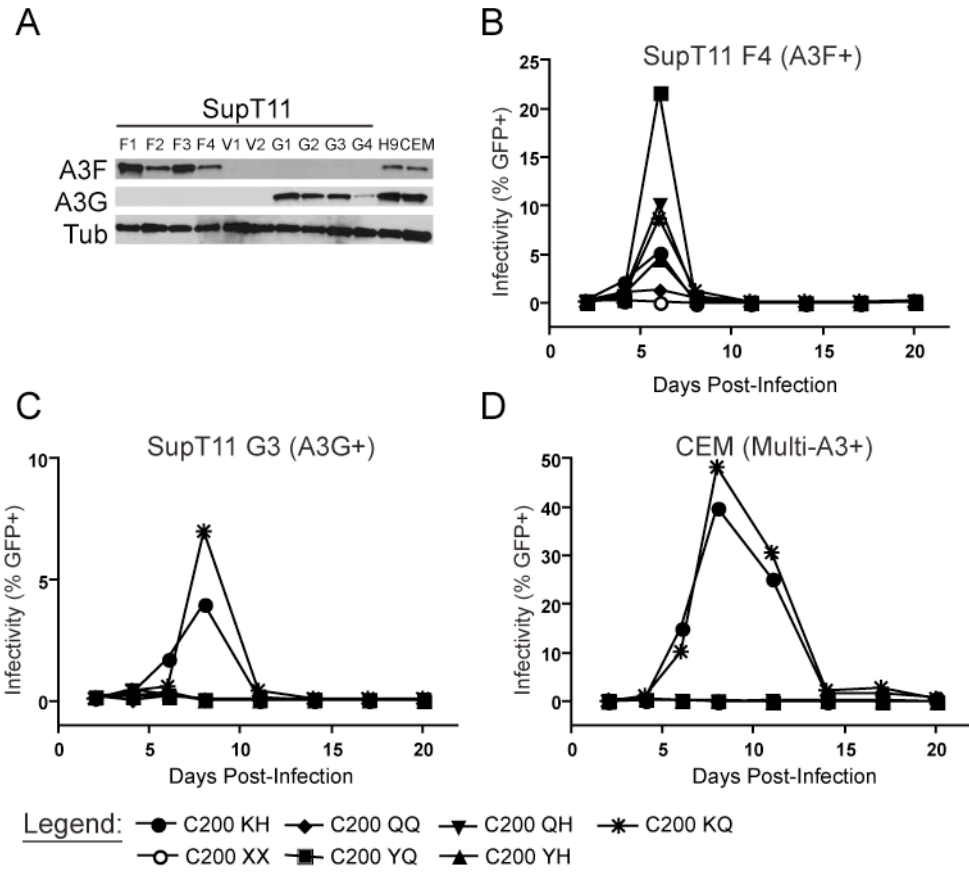
104

Figure 2-6: A3F-selected *vif* alleles are nonfunctional for the neutralization of A3G but can be rescued by restoration of the wildtype K26 residue.

Figure 2-6: A3F-selected *vif* alleles are nonfunctional for the neutralization of A3G but can be rescued by restoration of the wildtype K26 residue. (**A**) Western blots showing expression levels of A3F and A3G in the SupT11-derived cell lines used in these experiments as well as in H9 and CEM cells. Spreading infection curves from a starting MOI of 0.01 are shown for wildtype (C200 KH), Vif-deficient (C200 XX) and A3F-selected (C200 QQ and YQ) mutants as well as mutants completing the matrix of combinations of wildtype and selected residues at positions 26 and 27 of *vif* (C200 QH, YH and KQ) on SupT11 cells transfected with A3F (**B**) or A3G (**C**) as well as nonpermissive CEM cells (**D**). Results demonstrate that K26 is critical for the neutralization of A3G and the APOBEC3 repertoire found in CEM cells but not A3F. Results concordant with **2-6B-D** were observed using additional SupT11-derived cell lines expressing A3F or A3G as well as SupT11 clones transfected with a vector control and H9 cells (data not shown). Higher peaks than that shown in **2-6B** are usually observed with the C200 QQ virus (*e.g.* **Figures 2-1** and **2-3** and data not shown); in addition to those curves and many not shown, the ability of C200 QQ to spread in the presence of A3F is indicated by the fact that it efficiently kills the culture in which it replicates (see the legend to **Figure 2-1**).
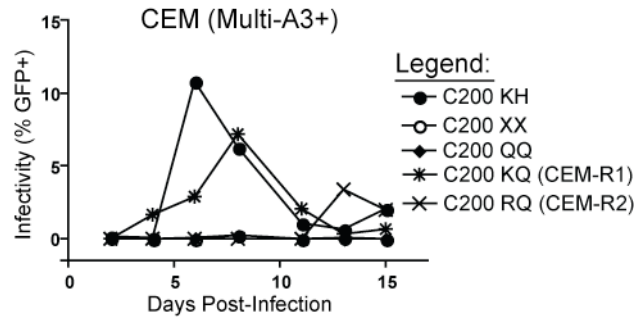
Figure 2-7: Long-term culture of A3F-resistant viruses in CEM cells selects for restoration of a positive charge at Vif residue 26. Passage of CEM-resistant viral isolates in CEM cells demonstrating their functional resistance to the nonpermissive phenotype of CEM. Phenotypic resistance of the selected alleles to the nonpermissive phenotype as encountered in CEM cells was also confirmed using viruses derived from molecular clones (**Table 2-1** and data not shown). A summary of the sequence evolution observed in resistant isolates that were confirmed by second passage and for which sequence was available is given in Experiment 5 of **Table 2-2**.

Postscript

The takeaway from **Chapter 2** is straightforward and neatly encapsulated in its title. That said, several points merit clarification. Only after completing this study did we understand that we had effectively stacked the deck against *vif* repair in our A3G selections since a lysine at position 26 is apparently the only option available for reversion in the presence of A3G (**Figure 2-6**). Efficacy of arginine may also be implied by the passages in naturally nonpermissive CEM cells (**Figure 2-7**), but R26-Q27 mutants were not clearly superior to Q26-Q27 in lowering the steady state levels of cotransfected A3G (data not shown) or in mediating spread in cell lines stably expressing A3G alone (Judd F. Hultquist, data not shown). These findings corroborate those of Dang *et al.*, where despite its conservative nature, the K26R mutation fails to restore Vif efficacy against A3G (Dang *et al.*, 2009). This may imply that K26 is a site of a post-translational modification that is somehow selectively critical to anti-A3G function, but we have not investigated this matter further. It is also possible that the K26R mutant may simply be a hypomorphic variant that more readily manifests in certain cells types than in others for any number of potential reasons.

Thus, while Vif anti-A3F function has a great deal of flexibility in reverting codons at positions 26 and 27, Vif anti-A3G function effectively has one, rare T-to-A transversion available to it. Despite the apparent contrast between the selective pressures applied by A3G and A3F, then, this phenomenon should not necessarily be taken as a direct comparison of the relative importance of A3F versus A3G for the maintenance of Vif function – but see **Chapters 5-6**.

Chapter 3: A Single Amino Acid in Human APOBEC3F Alters Susceptibility to HIV Vif

Foreword

This chapter is reproduced with permission from the following publication:

> Albin JS, LaRue RS, Weaver JA, Brown WL, Shindo K, Harjes E, Matsuo H and Harris RS. "A single amino acid in human APOBEC3F alters susceptibility to HIV Vif." *Journal of Biological Chemistry*. 2010; 285 (52): 40785-40792.

Following the work described in **Chapter 2**, we surmised that, despite the field's focus on A3G, further study of A3F may be more informative for our understanding of the APOBEC3-Vif interaction based simply on the fact that the genetic interaction between A3F and Vif was much clearer than that between A3G and Vif. Surprisingly, this has remained the case through several experimental iterations (see **Chapters 5-6**).

The natural next question, then, was to ask which portions of each protein are involved in this interaction. Since a highly influential alanine scanning mutagenesis paper had already identified putative A3F and A3G binding regions in Vif (but see **Chapter 5** and **Supplementary Chapter 2**), we opted to determine the Vif interaction site in A3F, which was only known in broad terms at this point based on chimera studies.

Summary

Human APOBEC3F (huA3F) potently restricts the infectivity of HIV in the absence of the viral accessory protein virion infectivity factor (Vif). Vif functions to preserve viral infectivity by triggering the degradation of huA3F but not rhesus macaque A3F (rhA3F). Here, we use a combination of deletions, chimeras and systematic mutagenesis between huA3F and rhA3F to identify E324 as a critical determinant of huA3F susceptibility to HIV Vif-mediated degradation. A structural model of the C-terminal deaminase domain of huA3F indicates that E324 is a surface residue within the α4 helix adjacent to residues corresponding to other known Vif susceptibility determinants in APOBEC3G and APOBEC3H. This structural clustering suggests that Vif may bind a conserved surface present in multiple APOBEC3 proteins.

Introduction

Human APOBEC3 proteins including APOBEC3F (huA3F) and APOBEC3G (huA3G) are DNA cytosine deaminases that restrict the infectivity of HIV in target cells following virion incorporation in producer cells [recently reviewed by (Albin, 2010; Goila-Gaur and Strebel, 2008; Henriet *et al.*, 2009)]. HIV overcomes this restriction activity by utilizing its accessory protein virion infectivity factor (Vif) to facilitate the degradation of APOBEC3 proteins in producer cells, thus preventing particle incorporation and restriction.

Previously, several groups identified specific changes in the N-terminal deaminase domain (NTD) of huA3G that affect the ability of HIV Vif to neutralize this restriction factor (Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004). The first of these studies sought to determine the basis for the observation that the Vif proteins of the lentiviruses infecting different species neutralize the A3G proteins of their natural host species but not the A3G proteins of other species (Mariani *et al.*, 2003). For example, African green monkey A3G (agmA3G) is susceptible to Vif from the simian immunodeficiency virus (SIV) that naturally infects *Chlorocebus aethiops* (agmSIV) but not to HIV Vif, while huA3G is susceptible to HIV Vif but not to agmSIV Vif. By substituting agmA3G residues into huA3G where the two differed, several groups identified D128 as a critical determinant of this species specificity (Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004). Subsequent mutational analyses have confirmed that huA3G D128 and surrounding residues including D130 impact HIV Vif-mediated degradation (Huthoff and Malim, 2007; Lavens *et al.*; Russell *et al.*, 2009b).

111

More recently, two reports showed that, in contrast with huA3G, huA3F is recognized at its C-terminal deaminase domain (CTD) by HIV Vif (Russell *et al.*, 2009b; Zhang *et al.*, 2008). One of these groups further narrowed the determinants of this recognition to amino acids 283-300, although individual amino acid changes critical for HIV Vif susceptibility were not identified in a manner analogous to the huA3G studies cited above. Thus, the residues of huA3F critical for the ability of HIV Vif to bind and degrade this restriction factor are presently unknown.

Here we identify a critical determinant of huA3F susceptibility to HIV Vif by comparing huA3F with the closely related but HIV Vif-resistant rhA3F (Virgen and Hatziioannou, 2007; Zennou and Bieniasz, 2006). Using chimeras between these orthologs as well as single-domain studies, we confirm that Vif recognizes the CTD of huA3F. Through systematic replacement of selected C-terminal huA3F residues with their corresponding rhA3F residues, we further identify huA3F QE323-324EK as a critical determinant of this differential susceptibility. Additional mutagenesis between these two residues revealed that mutation of E324 to the rhA3F lysine or to alanine results in resistance to HIV Vif-mediated degradation. To determine the three-dimensional context surrounding this residue, we created a model of the CTD of huA3F and found that E324 is a surface residue contained within the α4 helix that forms part of a broader surface shared with the linearly separate huA3F Vif interaction domain previously narrowed to residues 283-300 (Russell *et al.*, 2009b). Importantly, this analysis also revealed that the huA3F residues corresponding to three known Vif susceptibility determinants, D128 and D130 in huA3G and D/E121 in human APOBEC3H (huA3H), also cluster at this helix. These studies combine to suggest that a

112

conserved structural surface is targeted by HIV Vif en route to APOBEC3 neutralization

and degradation.

Results

HIV Vif recognizes the huA3F C-terminal deaminase domain – Two previous reports
have indicated that HIV Vif binds the CTD of huA3F (Russell *et al.*, 2009b; Zhang *et
al.*, 2008). To confirm and extend these results, we took two approaches. First, we created
chimeras between huA3F and rhA3F, which share 87% identity and 92% overall
similarity at the protein level (**Figure 3-1A**). This strategy takes advantage of the fact that
rhA3F is resistant to HIV Vif while huA3F is sensitive to HIV Vif (Virgen and
Hatziioannou, 2007; Zennou and Bieniasz, 2006). Therefore, by comparing the sensitivity
of chimeric proteins to HIV Vif-mediated degradation, one can broadly infer whether a
given chimera contains a site functionally recognized by HIV Vif. As shown in **Figure 3-
1B**, huA3F and the rhA3F/huA3F chimera containing the rhA3F NTD and the huA3F
CTD retained high sensitivity to HIV$_{IIIB}$ Vif-mediated degradation on cotransfection of a
given chimera with HIV Vif. In contrast, rhA3F and the huA3F/rhA3F chimera
containing the huA3F NTD and the rhA3F CTD were insensitive to the presence of HIV
Vif. To confirm these results, we cotransfected HA-tagged HIV Vif or SIV Vif with
GFP-tagged single deaminase domains of huA3F or rhA3F, respectively, and assessed
the stability of each domain. Both HIV Vif and SIV Vif destabilized the CTDs of huA3F
and rhA3F, respectively, while the corresponding NTDs were relatively unaffected
(**Figure 3-1C**). Thus, our data corroborate prior reports demonstrating that the A3F CTD
is necessary and sufficient for Vif-mediated degradation (Russell *et al.*, 2009b; Zhang *et
al.*, 2008).

A3F residues 323-324 affect the differential susceptibility of huA3F and rhA3F to HIV Vif – To more closely map the residues critical for functional neutralization of huA3F, we created a series of huA3F mutants containing 1-2 rhA3F residues at sites where these two differ within their CTDs. A schematic of the substitutions made is shown in **Figure 3-2A**. To test the Vif susceptibility of these mutants, we carried out single-cycle infectivity assays. As shown in **Figure 3-2B**, infectivity restoration upon the cotransfection of HIV$_{IIIB}$ Vif with most huA3F mutants was similar to that seen with wildtype huA3F. For the substitution QE323-324EK, however, Vif sensitivity was ablated (**Figure 3-2B**). These infectivity data correlated with producer cell huA3F levels, making this mutant phenotypically analogous to the control huA3G D128K.

Interestingly, neither of these substitutions is contained within the Vif binding region proposed by Russell *et al.* (Russell *et al.*, 2009b), residues 283-300, while the substitutions that do fall within this region (NLT298-300KLA) have no apparent phenotype [**Figure 3-2A-B** and Discussion]. Furthermore, huA3F D313H, which corresponds to the change D130K in the evolutionarily related NTD of huA3G, has no apparent phenotype [**Figure 3-2A-B**, **Discussion** and (25)].

To confirm the intrinsic Vif-resistance of huA3F QE323-324EK and eliminate the possibility that the C-terminal V5 tag initially used might affect our observations, we carried out single-cycle titration experiments using increasing levels of Vif cotransfected with a constant amount of huA3F, huA3F QE323-324EK, huA3G or huA3G DPD128-130KPK. As shown in **Figure 3-2C**, both huA3F QE323-324EK and huA3G DPD128-130KPK retained similar restriction regardless of Vif levels. The Vif-resistance of both

constructs was further confirmed by the intracellular stability of each in comparison with its wildtype control in the presence of Vif.

Because the single-cycle infectivity assays described to this point are vulnerable to potential overexpression artifacts [*e.g.* (Browne *et al.*, 2009; Miyagi *et al.*, 2007; Schumacher *et al.*, 2008), we also sought to assess the Vif-resistance of the huA3F QE323-324EK construct in a more physiologic setting. To that end, we created derivatives of a previously-described APOBEC3-deficient T cell line, SupT11, stably transfected with untagged huA3F QE323-324EK or huA3G DPD128-130KPK to go with our previously described derivatives expressing wildtype huA3F or huA3G (Albin *et al.*; Refsland *et al.*, 2010). The expression levels of huA3F and huA3G in each cell line used are shown in **Figure 3-2D**. We then initiated spreading infections at an MOI of 0.01 on these cell lines using Vif-deficient or wildtype IIIB or LAI-GFP viruses. All cell lines with the exception of the vector controls restricted the spread of Vif-deficient HIV_{IIIB} and $HIV_{LAI-GFP}$ (data not shown). In contrast, both wildtype HIV_{IIIB} and $HIV_{LAI-GFP}$, which differ at 20/192 Vif amino acids, spread efficiently on cell lines expressing wildtype huA3F or huA3G (**Figure 3-2E** and data not shown). Despite this, similar levels of huA3F QE323-324EK or huA3G DPD128-130KPK restricted the spread of even these Vif-proficient viruses (**Figure 3-2E** and data not shown). We thus conclude that huA3F QE323-324EK, like huA3G DPD128-130KPK, is fully resistant to HIV Vif and fully capable of inhibiting virus replication.

Reciprocal amino acid substitutions rhA3F EK323-324QE do not sensitize rhA3F to HIV Vif – To ask whether the reciprocal amino acid substitutions in rhA3F might render it

susceptible to HIV Vif in a manner analogous to the sensitization of agmA3G by the humanizing mutation K128D (Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004), we carried out single-cycle infectivity and expression analyses as above using rhA3F and rhA3F EK323-324QE. Under these conditions, rhA3F EK323-324QE showed no significant recovery in infectivity in the presence of HIV Vif over wildtype rhA3F (**Figure 3-3**). Thus, residues 323-324 are not exclusively responsible for the differential Vif sensitivity of rhA3F and huA3F. This observation is consistent with the emerging view that a larger surface on APOBEC3 proteins is recognized by Vif (See **Discussion**).

It is important to note that such separation of function experiments are not possible using SIV_{mac239} Vif, as this Vif neutralizes both huA3F and rhA3F (Zennou and Bieniasz, 2006). This parallels the results of several of the original papers characterizing huA3G D128K, which show that SIV_{mac239} Vif is able to neutralize both huA3G and rhA3G [*e.g.* (Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004)]. It is also consistent with our own studies suggesting that the Vifs of various species' lentiviruses are optimized for recognition of their own host species' APOBEC3Z3 proteins but often retain considerable activity against the APOBEC3Z3 proteins of other species (LaRue *et al.*, 2010).

Mutation of huA3F E324 alone alters functional susceptibility to HIV Vif in the absence of a quantitative reduction in physical binding – To further characterize the changes at residues 323-324, we created huA3F mutants with single cognate rhesus substitutions at each position as well as single and double alanine mutations at these positions and

assessed their restriction activities and Vif susceptibilities as before. This analysis revealed that any huA3F variant lacking glutamate at position 324 is resistant to HIV Vif regardless of the identity of residue 323, which correlates with intracellular A3F levels (**Figure 3-4A**). We therefore conclude that residue 324 is a single amino acid determinant of huA3F HIV Vif susceptibility.

To determine whether mutation of E324 alters the quantitative binding of Vif to huA3F, we cotransfected HA-tagged Vif with V5-tagged huA3F and variants mutated at residues 323 and 324. We then immunoprecipitated HA-tagged Vif from these lysates and blotted for associated V5-tagged huA3F. Despite the resistance of huA3F E324 variants to HIV Vif-mediated degradation, however, we found that mutation of this residue does not reduce coimmunoprecipitation with HIV Vif relative to wildtype (**Figure 3-4B-C**). These results were confirmed by reciprocal immunoprecipitation of the V5-tagged huA3F or huA3F E324K and blotting for cotransfected, untagged Vif in both RIPA and NP40 lysis buffers (data not shown).

Determinants of HIV Vif recognition localize to the α4 helix of a susceptible deaminase domain – To visualize E324 in its three-dimensional context, we created a model of the CTD of huA3F based on a recent crystal structure of the huA3G CTD [PDB ID 3IR2, (Shandilya *et al.*, 2010). As shown in **Figure 3-5A-B**, E324 is located on the surface of the α4 helix. Importantly, this region is also adjacent to the linearly separate stretch of amino acids previously implicated in the huA3F interaction with HIV Vif, residues 283-300 (Russell *et al.*, 2009b). We also noted, however, that several additional negatively charged residues occurred on or near the surface of this helix: D311, D313 and E316. On

aligning these residues to those found in other Vif-susceptible deaminase domains in **Figure 3-5C**, we noted that they each align to a previously-described APOBEC3 determinant of Vif susceptibility (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Lavens *et al.*; Mangeat *et al.*, 2004; Russell *et al.*, 2009b; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004). D311 and D313 in huA3F correspond to D128 and D130 in the NTD of huA3G, while huA3F E316 corresponds to E/D121 in human APOBEC3H (huA3H). Thus, all known APOBEC3 determinants of Vif susceptibility cluster along the surface of the α4 helix, and all are negatively charged.

Discussion

The studies described here are the first to identify a single amino acid determinant of the susceptibility of huA3F to HIV Vif. This represents an important advance in our understanding of the HIV Vif-huA3F interaction, the relevance of which is strongly supported by a large body of work demonstrating the potency of huA3F-mediated restriction of HIV [*e.g.* (Bishop *et al.*, 2004; Liddament *et al.*, 2004; Wiegand *et al.*, 2004; Zheng *et al.*, 2004)]. Our own long-term viral evolution studies have also suggested that functional neutralization of huA3F by HIV Vif is required for the virus propagate in the presence of huA3F (Albin *et al.*, 2010a). Thus, shielding the α4 region of huA3F described here from HIV Vif may represent a viable strategy for the development of novel pharmacotherapies for HIV infection [*e.g.* (Albin, 2010; Harris and Liddament, 2004)].

Our work confirms prior reports that broadly localized Vif interaction to the CTD of huA3F [(Russell *et al.*, 2009b; Zhang *et al.*, 2008), **Figure 3-1**]. An additional recent report on the existence of a Vif-susceptible splice variant of huA3F composed largely of the CTD is also consistent with these data (Lassen *et al.*, 2010).

Our identification of a single amino acid determinant of HIV Vif susceptibility in huA3F echoes several prior reports localizing Vif susceptibility in huA3G and huA3H in that the residue identified is a single negative charge localized to the surface of an APOBEC3 protein [(Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004), **Figures 3-4** and **3-5**]. E324 differs from these reports in one key respect, however, as a charge substitution was involved in all prior reports. For example, huA3G D128A has no phenotype, while D128K is Vif-resistant [*e.g.* (Bogerd *et al.*,

120

2004; Huthoff and Malim, 2007; Schröfelbauer *et al.*, 2004)]. The fact that both alanine and lysine substitutions at huA3F E324 ablate Vif susceptibility (**Figure 3-4A**) suggests that E324 is required either for overall stability of the broader Vif binding surface or for direct functional interaction with HIV Vif. Changes to this residue do not, however, affect the ability of huA3F to restrict HIV; in fact, none of the changes described in these studies affected restriction activity (*e.g.* **Figures 3-2**, **3-3**, **3-4**).

It is notable that E324 does not fall within the region previously found by Russell *et al.* to be critical for HIV Vif recognition of huA3F (Russell *et al.*, 2009b). These residues, 283-300, encompass most of the α3 helix, which is structurally adjacent to α4 and E324 and appears to form a common surface [**Figure 3-5A-B** and (Russell *et al.*, 2009b)]. It is therefore possible that E324 may cooperate with residues in α3 and/or α4 to create a stable surface recognized by HIV Vif, in which case mutational alteration of any critical component of this putative Vif interaction node may affect Vif sensitivity. Alternatively, Russell *et al.* used chimeras between huA3F and huA3G to map the huA3F Vif interacting region (Russell *et al.*, 2009b). The C-terminal deaminase domains of huA3F and huA3G, however, are evolutionarily divergent Z2 and Z1 types, respectively (25). This means that chimeras in this domain will contain a relatively large number of amino acid substitutions versus wildtype. For example, only half of residues 283-300 are biochemically similar or identical between huA3F and huA3G. Thus, we think it likely that the structure of huA3F-huA3G chimeras in this region will be altered relative to huA3F, which may affect interaction with Vif.

In addition to suggesting the structural unity of our findings with those of Russell *et al.*, our model of the huA3F CTD allowed us to make an important observation about

the nature of the region surrounding E324. Namely, both the α4 and the neighboring α3 helix have a number of negatively charged surface residues (**Figure 3-5A-B**). This led us to align these surface residues with those in other Vif-susceptible APOBEC3 deaminase domains such as the huA3G NTD and huA3H, which showed that each negatively-charged surface residue in the α4 helix of huA3F corresponds to a known negatively charged determinant of Vif susceptibility in another APOBEC3 protein (**Figure 3-5**). Thus, while determining the identity of all the amino acid residues with which Vif interacts (*i.e.* the broader Vif binding surface in APOBEC3 proteins) will require a great deal of future genetic and structural study, it is intriguing that all known Vif susceptibility determinants map to the same structural motif. This implies a degree of structural conservation among APOBEC3-Vif interaction surfaces that would not be apparent from a simple linear comparison of these single amino acid determinants.

While the lack of functional interaction between Vif and huA3F E324 variants is clear, this appears to be due to a qualitative change in the nature of the Vif-huA3F interaction in E324 mutants since coimmunoprecipitation of huA3F E324 mutants with HIV Vif is unimpaired. Our data therefore support the potential for both qualitative and quantitative changes in huA3F binding to Vif that may alter susceptibility. In the absence of structural data, however, we are unable to explain the nature of the qualitative defect found in huA3F E324 mutants. It is possible that the qualitative defect in E324 mutants may involve a conformational change that prevents productive interaction with HIV Vif. Alternatively, mutation of E324 may impair the recruitment of other components of the E3 ligase complex by HIV Vif en route to degradation. It is also conceivable that E324 mutants may have a functionally relevant, altered affinity for HIV Vif that is not readily

apparent by coimmunoprecipitation.

In summary, we have described here a single amino acid determinant of huA3F susceptibility to HIV Vif. This important advance in our understanding of the Vif-huA3F interaction echoes the single amino acid determinants previously identified in other APOBEC3 proteins, as all are negatively charged residues that may interact directly with the highly basic Vif protein. Importantly, the observation that all of these single amino acid determinants cluster along the α4 helix raises the exciting possibility that certain features of the APOBEC3-Vif interaction may be structurally conserved, which would facilitate the design of hypothetical single molecules which may simultaneously block the functional interaction of Vif with multiple APOBEC3 proteins. Indeed, while its exact mechanism of action remains unknown, the lead compound RN-18 provides proof of concept for just such a scenario (Nathans *et al.*, 2008).

Materials and Methods

Plasmid DNA construction and site-directed mutagenesis – All constructs were confirmed by DNA sequencing. huA3F and huA3G coding sequences correspond to those found in GenBank NM_145298 and NM_021822, respectively. rhA3F was provided by Dr. Theodora Hatziioannou (Aaron Diamond AIDS Research Center, New York) (Virgen and Hatziioannou, 2007; Zennou and Bieniasz, 2006). Substitutions of rhA3F residues into huA3F were based on alignment between huA3F and rhA3F reference sequences NM_145298 and NM_001042373.1. pcDNA3.1-V5, -huA3F-V5 and -huA3G-V5 have been described, and pcDNA3.1-huA3G-V5 D128K was similarly derived (Albin *et al.*, 2010a).   A3F domain chimeras were made using overlapping PCR (LaRue *et al.*, 2008). PCR products were digested with KpnI/XhoI and ligated into similarly cut pcDNA3.1-V5. Single-domains of huA3F and rhA3F were amplified using primers containing SacI/SalI sites and cloned into similarly cut pEGFP-N3 (Clontech). Full-length huA3F-GFP has been described (Stenglein and Harris, 2006); NTD = residues 1-191; CTD = residues 192-373. huA3F PE281-282LD, NLT298-300KLA, T303A and D313H were introduced into the pcDNA3.1-huA3F-3xHA construct (Stenglein and Harris, 2006) by site-directed mutagenesis using Pfu polymerases (Stratagene). The 3xHA tag was subsequently replaced with a V5 tag (Albin *et al.*, 2010a). All other mutations were introduced directly into pcDNA3.1-huA3F-V5.

HIV$_{IIIB}$ and SIV$_{mac239}$ Vif as well as a vector derived from pVR1012 have been described (Albin *et al.*, 2010a; LaRue, 2010). An untagged, codon-optimized version of HIV$_{IIIB}$ Vif was made by PCR amplification and ligation of the coding region into the SalI/BamHI segment of the original construct. The codon-optimized translated Vif open

124

reading frames are those of HIV$_{IIIB}$ (GenBank EU541617) and SIV$_{mac239}$ (GenBank AY588946).

Proviral plasmid HIV$_{IIIB}$ is a nucleotide A200C derivative of pIIIB (Albin *et al.*, 2010a). A Vif-deficient A200C HIV$_{IIIB}$ derivative containing a previously described deletion in *vif* made by overlap extension PCR was used in spreading infections (Gibbs *et al.*, 1994). A Vif-deficient HIV$_{IIIB}$ derivative containing tandem stop codons at positions 26-27 of *vif* was used for all single-cycle infectivity experiments and has been previously described (Albin *et al.*, 2010; Haché *et al.*, 2008). Wildtype and Vif-deficient HIV$_{LAI-GFP}$ were kindly provided by Dr. Mario Stevenson (University of Massachusetts, Worcester, MA).

Cell lines – 293T cells were maintained in DMEM supplemented with 10% fetal bovine serum and, in some cases, penicillin/streptomycin. CEM-GFP reporter cells were maintained in RPMI supplemented with 10% fetal bovine serum, penicillin/streptomycin and β-mercaptoethanol (Gervaix *et al.*, 1997; Haché *et al.*, 2008).

Stability of A3F chimeras in the presence of HIV and SIV Vifs – At 50% confluency in 6-well plates, 293T cells were transfected using Trans-IT transfection reagent (Mirus Bio) with 100 ng A3-V5 and 25 ng Vif-HA. After 48 hours, cell lysates were harvested and resuspended in 2x sample buffer 25 mM Tris pH 6.8, 8% glycerol, 0.8% SDS, 2% β-mercaptoethanol, 0.02% bromophenol blue, boiled for 10 minutes, and run on a 12% SDS-PAGE gel prior to transfer to a PVDF membrane (Millipore). Membranes were probed with mouse anti-V5 (Invitrogen), mouse anti-HA.11 (Covance) or mouse anti-α-

tubulin (Covance) primary antibodies followed by incubation with horseradish peroxidase (HRP)-conjugated goat anti-mouse antibodies. Membranes were developed using HyGLO chemiluminescent HRP detection reagent (Denville Scientific) and exposed to film. Blots were stripped using 0.2 M glycine, 1.0% SDS, 1.0% Tween-20, pH 2.2 between sequential probing with primary antibodies.

Single-cycle infectivity assays – 250,000 293T cells were plated in 2 mL DMEM in 6-well plates. One day later, Trans-IT transfection reagent (Mirus Bio) was used to cotransfect these cells with 1.6 μg Vif-deficient HIV_{IIIB}, 100 ng of a codon-optimized HIV_{IIIB} Vif-HA expression construct (or 50-200 ng supplemented to 200 ng total with a vector control in **Figure 3-2C**) and 200 ng of a given APOBEC3-V5 construct. Approximately two days later, virus-containing supernatants were filtered through 0.45 μm PVDF filters (Millipore) and 75 or 150 μL were used to infect 25,000 CEM-GFP reporter cells plated at a final total volume of 250 μL. At this time, 293T cells were resuspended with 1 mL phosphate buffered saline (PBS) and 500 μL were spun down and resuspended in 250 μL lysis buffer [25 mM HEPES pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 50 μM $ZnCl_2$, 10% glycerol and 1% Triton X-100 supplemented with 50 μM MG132 (American Peptide) and complete protease inhibitor (Roche)] for analysis of APOBEC3 intracellular stability. Three days post-harvest, CEM-GFP reporter cells were fixed with 4% paraformaldehyde and analyzed by flow cytometry on a Beckman-Coulter Quanta MPL or a Becton-Dickinson LSR II to determine infectivity as measured by the percentage of GFP-positive cells. Relative infectivity was calculated by normalizing the infectivity of viruses produced in the presence of each APOBEC3 protein +/- Vif to the

infectivity of viruses produced in the presence of an APOBEC3 vector control +/- Vif in each experiment. Data shown represent the mean and SEM of the number of independent transfection-infection series indicated in the figure legend.

Analysis of intracellular APOBEC3 stability – Cotransfected cells were lysed as described above, and aliquots were mixed with a 5x or 7.5x version of the sample buffer described to a final concentration of 2x and boiled for 10 minutes. Proteins were then separated by 10% SDS-PAGE and transferred to a PVDF membrane. Membranes were probed as described above and stripped using 62.5 mM Tris pH 6.8, 2% SDS and 100 mM β-mercaptoethanol at 50°C prior to sequential blocking and reprobing.

Analysis of intracellular APOBEC3 expression – APOBEC3-expressing and vector control derivatives of SupT11 as well as the model nonpermissive cell lines CEM and H9 were grown to confluence in 10 cm dishes. $5x10^6$ cells were then lysed in 250 µl and analyzed for huA3F or huA3G expression using antibodies #11474 or #10201 from Drs. Michael Malim or Jaisri Lingappa, respectively, obtained through the AIDS Research and Reference Reagent Program.

Virus titration and spreading infections – Viruses were produced by plating $3.5x10^6$ 293T cells in 10 cm dishes and one day later transfecting those cells with 5-10 µg of a given proviral plasmid using Trans-IT transfection reagent (Mirus Bio). Approximately two days post-transfection, supernatants were harvested, and different volumes were used to infect 150,000 CEM-GFP reporter cells at a constant total volume of 1 mL in 24-well

plates. Three days later, these CEM-GFP cells were fixed in 4% paraformaldehyde, and the total percentage of GFP-positive cells was quantified by flow cytometry as before. Linear regression was then employed to determine the volume of a given viral stock required to initiate infection at a CEM-GFP MOI of 0.01.

Spreading infections were initiated by infecting 150,000 cells of a given cell line at an MOI of 0.01 in a total volume of 1 mL in 24-well plates. Cultures were subsequently split and fed as necessary to prevent cell overgrowth. Viral spread was monitored by periodically harvesting 150 µL of supernatant from each culture and using it to infect 25,000 CEM-GFP cells at a final volume of 250 µL in 96-well plates. Three days post-infection, these CEM-GFP were fixed in 4% paraformaldehyde, and the percentage of GFP-positive cells was analyzed by flow cytometry on a Becton-Dickinson LSR II.

Homology Modeling of the huA3F C-terminal Deaminase Domain – The A3F186-373 model was generated using YASARA (Krieger *et al.*, 2002) based on the crystal structure of A3G191-384 2K3A [PDB ID 3IR2, (Shandilya *et al.*, 2010)]. Alignment with the A3G191-384 sequence (**Figure 3-6**) was iteratively optimized using related SwissProt and TrEMBL sequences, the predicted secondary structure and the structural information of the template. Knowledge-based and electrostatic interactions in unrestrained molecular dynamics with explicit solvent molecules were used to refine amino acid side chain geometry. Insertions were accounted for by a search of the Protein Data Bank for loop ends superimposable with model anchor points. Further optimization was achieved by placement of loops into their lowest energy conformations.

Coimmunoprecipitation – Coimmunoprecipitation of V5-tagged huA3F with HA-tagged HIV Vif or with HA-tagged Vif with mutation of the conserved BC Box residues SLQ>AAA that ablate Vif-mediated degradation was carried out by lysing cotransfected 293T cells with RIPA buffer [50 mM Tris pH 8.0, 150 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS supplemented with complete protease inhibitor (Roche)]. Lysates were then incubated with 2.5 µL of mouse anti-HA.11 at 4°C followed by the addition of 40 µL Dynabeads Protein G (Invitrogen). Immunoprecipitated complexes were isolated by magnetic separation, washed four times with PBS and eluted by addition of 30 µL of 5x sample buffer as above. SDS-PAGE and Western blotting were then carried out as before. Western blots were quantified by analysis with ImageJ software (Rasband, 1997-2009). Binding quantification represents the intensity of V5 bands immunoprecipitated divided by the intensity of the corresponding lysate bands and normalized to the ratio found for huA3F in a given experiment.
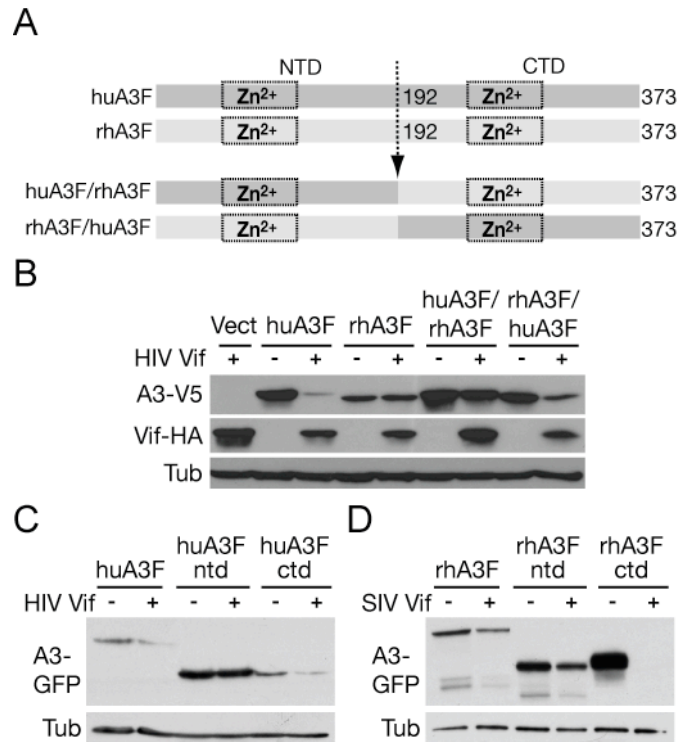
Figure 3-1: Susceptibility of huA3F to HIV Vif maps to the huA3F C-terminal deaminase domain. (A) A schematic depiction of the chimeras used in **Figure 3-1B**. **(B)** A cotransfection experiment demonstrating the instability in the presence of HIV Vif of chimeras between rhA3F and huA3F that contain the huA3F CTD. **(C-D)** Cotransfection experiments demonstrating that the CTDs of huA3F and rhA3F are destabilized by HIV Vif or SIV Vif, respectively, while the corresponding NTDs remains highly expressed.
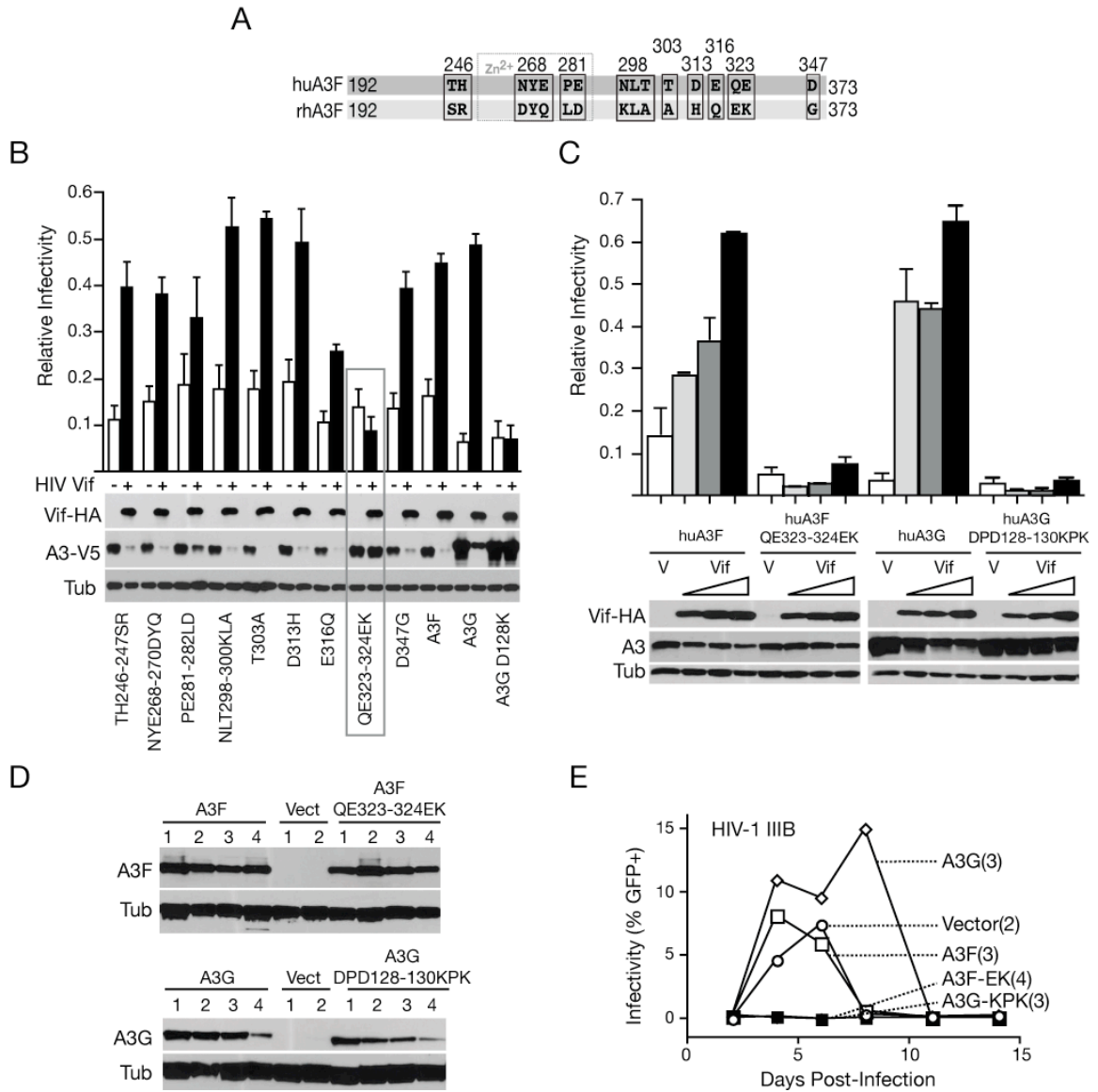
Figure 3-2: Substitution of rhA3F residues at positions 323-324 of huA3F results in phenotypic Vif resistance.

Figure 3-2: Substitution of rhA3F residues at positions 323-324 of huA3F results in phenotypic Vif resistance. (A) A schematic depiction of the CTDs of huA3F and rhA3F with key changes tested through site-directed mutagenesis for their effects on the Vif-sensitivity of huA3F indicated as numbered from the first position changed in a given set. **(B)** Single-cycle infectivity data quantifying the effects of HIV Vif on the rescue of HIV infectivity. All mutants remained competent for restriction. Relative infectivity represents the mean and SEM of four independent experiments. Western blots demonstrating intracellular APOBEC3 and Vif expression levels are shown below blotting for V5 or HA, respectively, and are derived from the producer cells of one of these four experiments. **(C)** Single-cycle infectivity data demonstrating the continued restriction of HIV by untagged huA3F QE323-324EK and huA3G DPD128-130KPK in the presence of increasing amounts of Vif. Relative infectivity represents the mean and SEM of two independent experiments done in duplicate. The Western blots demonstrating the stability of Vif-resistant huA3F and huA3G variants in the presence of Vif correspond to one of these infectivity experiments. **(D)** Western blots showing the expression levels of huA3F, huA3G and their Vif-resistant variants in the cell lines used. **(E)** Spreading infection curves demonstrating that wildtype HIV_{IIIB} is restricted by Vif-resistant variants huA3F QE323-324EK and huA3G DPD128-130KPK but not the corresponding wildtype proteins. Similar results were obtained with three additional cell lines stably transfected with A3F, A3G or the Vif-resistant variants thereof as well as one additional vector control cell line (data not shown). Open symbols indicate vector control or huA3F- or huA3G-expressing cell lines. Closed symbols indicate cell lines expressing huA3F

QE323-324EK or huA3G DPD128-130KPK. The y-axis is offset from zero to permit visualization of curves yielding no detectable spread.
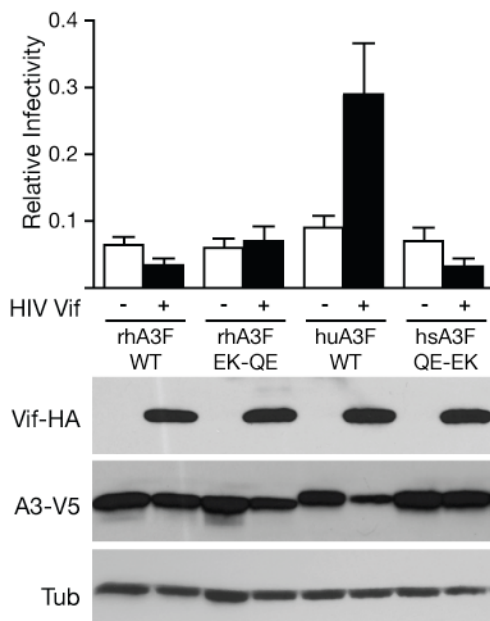
Figure 3-3: Substitution of human residues at positions 323-324 of rhA3F does not sensitize rhA3F to HIV Vif. Single-cycle infectivity experiments demonstrating that substituting the human residues at positions 323-324 of rhA3F does not sensitize this restriction factor to permit infectivity recovery in the presence of Vif. Data represent the mean and SEM of three independent experiments. Western blots corresponding to one of the single-cycle experiments shown demonstrate the correlation between intracellular stability of APOBEC3 variants and functional recovery in infectivity.
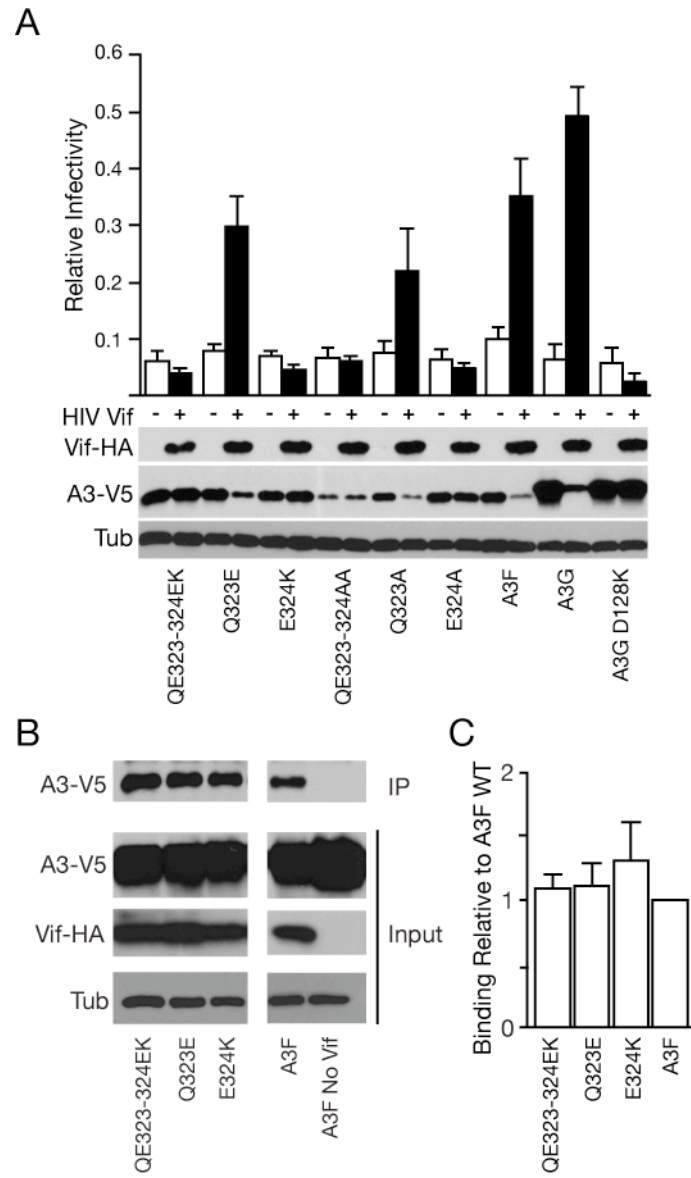
Figure 3-4: The identity of residue 324 is a primary determinant of the degradation sensitivity of huA3F to HIV Vif.

Figure 3-4: The identity of residue 324 is a primary determinant of the degradation sensitivity of huA3F to HIV Vif. (A) Single-cycle infectivity data quantifying the restriction and Vif-sensitivity phenotypes of single and double human-to-rhesus and human-to-alanine mutations at positions 323 and/or 324 of huA3F, where only mutations at position 324 ablate Vif-responsiveness. Data represent the mean and SEM of three independent experiments. Western blots corresponding to one of these experiments demonstrating that mutations at position 324 of huA3F result in resistance to Vif-mediated degradation are shown below. **(B)** A representative experiment demonstrating the lack of effect of mutations at huA3F positions 323-324 on coimmunoprecipitation with HA-tagged Vif. The images shown are all taken from the same blot where the A3F and A3F No Vif lanes have been digitally juxtaposed with the experimental mutants. **(C)** Quantification of the results from a total of five independent experiments similar to and including that shown in **Figure 3-4B**. Relative binding represents the ratio of IP V5 signal to cellular V5 signal normalized to the ratio found for A3F in a given experiment, which is set to 1.0.

A

B
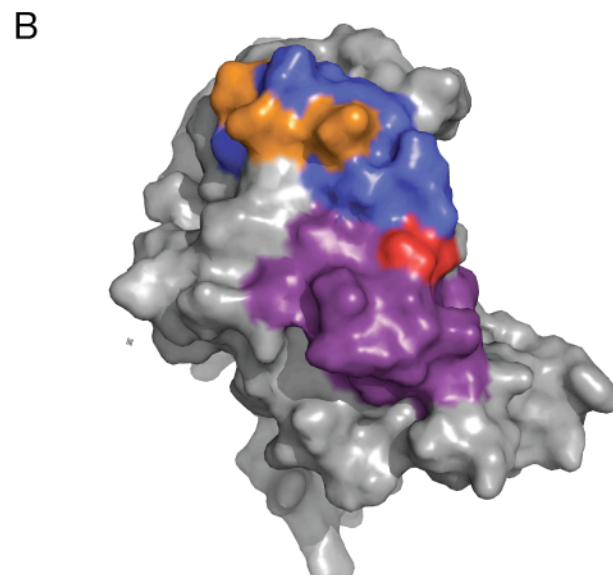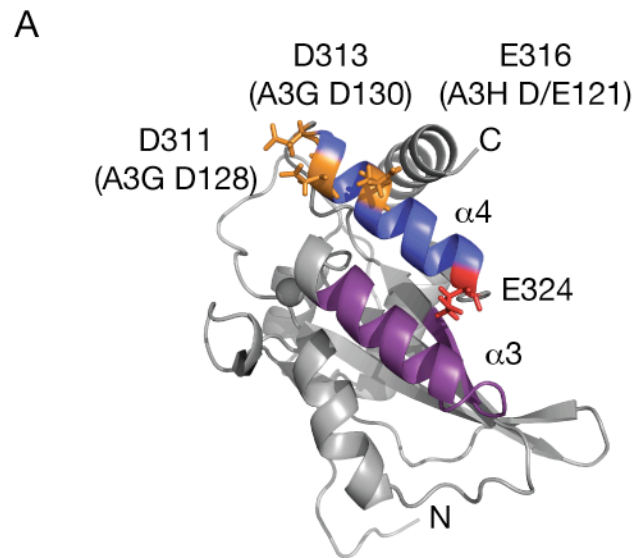
C

```
huA3F 311-324  DTDYQEGLRSLSQ--E
huA3G 128-143  DPDYQEALRSLCQKRD
huA3H 116-128  CKPQQDGLRLLCG---
```

Figure 3-5: A model structure of the C-terminal deaminase domain of huA3F.

Figure 3-5: A model structure of the C-terminal deaminase domain of huA3F. (A) A ribbon diagram depicting the CTD of huA3F. The region encompassing the huA3F equivalents of all known single amino acid determinants of Vif-sensitivity is shown in blue, with D311, D313 and E316 shown in orange and E324 in red. The region previously implicated in huA3F interaction with Vif (residues 283-300) is colored purple. **(B)** The predicted surface of the huA3F CTD. **(C)** An alignment of residues in the α4 helix encompassing known determinants of Vif susceptibility.
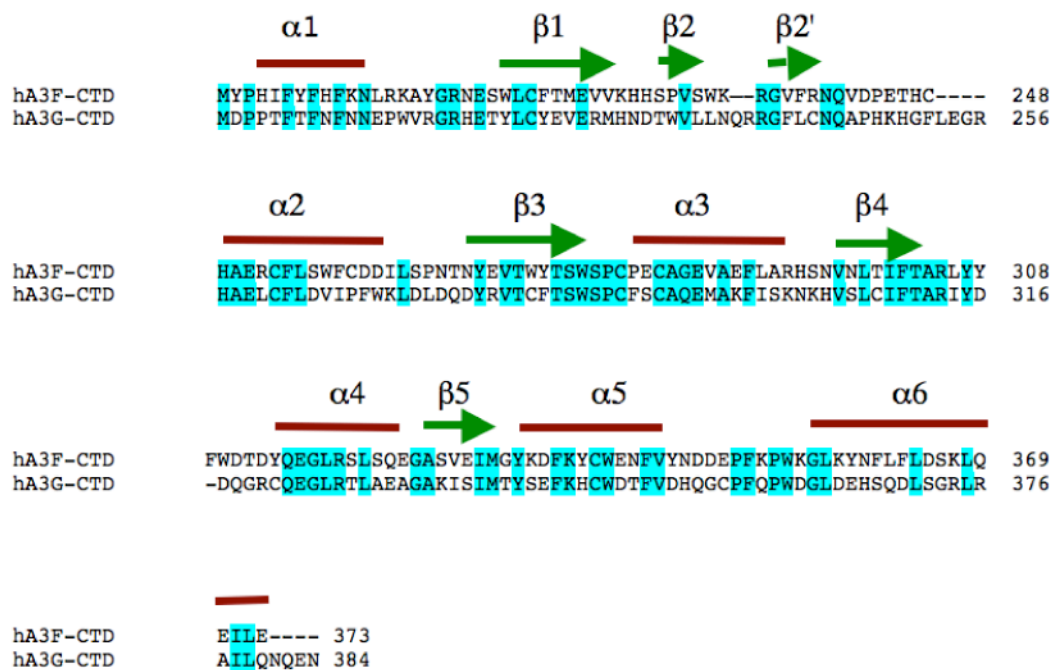
Figure 3-6: Alignment of huA3G and huA3F used for structural modeling. An alignment of huA3G and huA3F including secondary structural elements used in making the model in **Figure 3-5A-B** is shown.

Postscript

The study described in **Chapter 3** appears modest in its scope, a single residue story using a standard toolbox. It is my opinion, however, that **Figure 3-5** represents an important insight that will resonate throughout the field as we come to understand more about the APOBEC3-Vif interaction, starting with **Chapter 5** and someday including co-crystal structures of complex components to complement the genetic and functional data. This chapter yielded the tool and the insight necessary to start to bring some sense of order to the otherwise complicated mutagenesis data on APOBEC3-Vif interactions (compare **Table 1-2** with **Chapter 5**).

Chapter 4: Dispersed Sites of HIV Vif-dependent Polyubiquitination in the DNA

Deaminase APOBEC3F

Foreword

The data presented in this chapter derive from the following manuscript *in revision*.

> Albin JS, Anderson JS, Johnson JR, Harjes E, Matsuo H, Krogan NJ and Harris
> RS. "Dispersed sites HIV Vif-dependent polyubiquitination in the DNA
> deaminase APOBEC3F." *in revision*

Chapter 3 described the genetic identification of the Vif interaction site in A3F. Aside

from a Vif interaction site, there must also be ubiquitin acceptor sites in A3F in order for

Vif to target the deaminase for proteasomal degradation. A previous study using A3G had

suggested that these could be clearly narrowed to a set of four out of the 20 internal

lysines in A3G (Iwatani *et al.*, 2009). Based on the structural clustering of these targeted

lysines, these authors proposed a novel model of APOBEC3-Vif interactions in which

binding of Vif to one domain in a double-domain deaminase such as A3F or A3G orients

the opposite domain for polyubiquitination. We therefore set out to determine whether

polyubiquitinated lysines in A3F might be similarly identified and whether the Vif

interaction model previously proposed would prove valid for A3F.

Summary

APOBEC3F (A3F) and APOBEC3G (A3G) are DNA cytosine deaminases that potently restrict human immunodeficiency virus type 1 (HIV) replication when the virus is deprived of its accessory protein virion infectivity factor (Vif). Vif counteracts these restriction factors by recruiting A3F and A3G to an E3 ubiquitin ligase complex that stimulates their polybiquitination and proteasomal degradation. While previous efforts have identified a Vif-dependent recognition site in APOBEC3 proteins centered on the α4 helix of Vif-susceptible deaminases, less is known about the downstream ubiquitin acceptor sites that are targeted. A prior report aimed at identifying polyubiquitinated residues in A3G, however, proposed an antiparallel model of A3G interaction with the Vif-E3 ubiquitin ligase complex wherein Vif binding at one terminus of A3G orients the opposite terminus for polyubiquitination (Iwatani *et al.*, 2009). To test the generalizability of this model, we carried out a complete mutagenesis of the lysine residues in A3F and used a complementary, unbiased proteomic approach to identify ubiquitin acceptor sites targeted by Vif. Our data indicate that internal lysines are the dominant ubiquitin acceptor sites in both A3F and A3G, but in contrast with the proposed antiparallel model, we find that the Vif-dependent polyubiquitination of A3F and A3G can occur at multiple acceptor sites dispersed along lysine-enriched surfaces of both the N- and C-terminal deaminase domains, suggesting alternative models for binding of APOBEC3 proteins to the Vif-E3 ubiquitin ligase complex and diminishing enthusiasm for the amenability of APOBEC3 ubiquitin acceptor sites to therapeutic intervention.

Introduction

Human A3F and A3G are DNA cytosine deaminases capable of inhibiting HIV reverse transcription and integration, most prominently through their active C-terminal deaminase domains, which introduce massive levels of G-to-A mutations in the nascent provirus that contribute to incomplete reverse transcription and render hypermutated genomes hypofunctional [reviewed (Albin, 2010; Chiu and Greene, 2008; Malim, 2009)]. To achieve this restriction activity, A3F and A3G must encapsidate during viral budding in producer cells, but HIV typically circumvents this inhibitory APOBEC3 activity by utilizing its accessory protein Vif as a substrate receptor to link A3F and A3G to an E3 ubiquitin (Ub) ligase complex consisting of CBFβ, ELOB and ELOC, CUL5 and RBX2, which results in the polyubiquitination (polyUb) and subsequent proteasomal degradation of these restriction factors [see recent papers (Jäger *et al.*, 2011; Zhang *et al.*, 2011) and references therein].

Much effort has been devoted to identifying the determinants of the APOBEC3-Vif interaction critical for this degradative process, including a variety of changes in Vif that result in the functional inactivation of its anti-APOBEC3 activity [*e.g.* (Russell and Pathak, 2007; Simon *et al.*, 2005); reviewed in (Albin, 2010; Smith *et al.*, 2009)]. On the APOBEC3 side of this direct, host-pathogen interaction, initial efforts suggested that the Vif binding sites in A3F and A3G were structurally distinct, occurring in the C-terminus of the former and in the N-terminus of the latter at unrelated residues (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Lavens *et al.*, 2010; Mangeat *et al.*, 2004; Russell *et al.*, 2009b; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004). However, independent studies indicated that these single amino acid determinants of Vif recognition typically occur

143

within a common region within a given susceptible APOBEC3 deaminase domain centered on the α4 helix, suggesting that a conserved structural determinant may be targeted by Vif [DPD128-130 in A3G (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004); E324 and E289 in A3F (Albin *et al.*, 2010b; Smith and Pathak, 2010) and D/E121 in A3H (Zhen *et al.*, 2010).

Less is known about the downstream Ub acceptor sites targeted for polyUb during the degradation process. Iwatani *et al.* have previously reported that four lysines in the C-terminus of A3G are the Ub acceptor sites required for the Vif-mediated degradation of this enzyme (Iwatani *et al.*, 2009). Combined with the aforementioned studies localizing the Vif interaction to the N-terminal half of the A3G, this led to the proposal of an antiparallel model of APOBEC3 binding to the Vif-E3 Ub ligase complex wherein Vif interacts with one deaminase domain, thereby orienting the second domain for polyUb by an activated E2~Ub conjugate at the opposite end of the CUL5 scaffold.

To test this model and potentially enhance our understanding both of APOBEC3 binding to the Vif-E3 Ub ligase complex and of how these substrate acceptor sites might be utilized to block the degradation of APOBEC3 proteins, we set out to define the sites of polyUb in A3F. Consistent with results previously reported by Iwatani *et al.*, we find that internal lysines are the dominant Ub acceptor sites in both A3F and A3G. Analysis of the specific residues available for functional polyUb, however, reveals that these are dispersed throughout both domains of A3F at lysine residues clustered on one side of model structures of the A3F N-terminal and C-terminal deaminase domains, opposite the Vif interaction site in the latter case. Furthermore, mutation of the lysine residues determined by Iwatani *et al.* to be the sites of polyUb in A3G confers only partial Vif

144

resistance. Consistent with these genetic data, mass spectrometric analysis reveals Vif-dependent Ub modification of at least 6 sites in A3F and 10 sites in A3G localized to both the N- and C-terminal deaminase domains of each. We conclude, therefore, that the lysine residues available for Vif-dependent polyUb in human APOBEC3 proteins are diverse and unlikely to be leveraged for novel therapeutics. To explain this flexibility, we further propose an alternative model of APOBEC3 binding to the Vif-E3 Ub ligase complex.

Results

The Sites of Vif-mediated polyUb in A3F are Distributed Throughout the Protein – The lysine codons in A3F cluster into three linear groups separated by unique EcoRI and BamHI sites. We first used serial mutagenesis to convert each of these groups from K-to-R. Next, we joined the three regions together to make the panel of K-to-R mutants including a derivative completely devoid of lysines, A3F-19KR, and tested the restriction activity and Vif susceptibility of the resultant proteins in a single cycle of replication (**Figure 4-1A**).

The restriction activity of all A3F mutants was similar to that of the wildtype protein, and the K-to-R mutation of any single linear region resulted in A3F variants with Vif susceptibility similar to that of the wildtype protein as evidenced both by the recovery of infectivity and by the decrease in steady state A3F levels in the presence of Vif (**Figure 4-1B**). Similarly, the combination of any two linear regions of K-to-R mutants resulted in somewhat more resistant but still notably Vif-susceptible variants of A3F. Only the variant devoid of all lysines, A3F-19KR, was largely resistant to the effects of Vif. Taken together, these data indicate that at least one lysine residue in each linear region is available for Vif-mediated polyUb. We also noted that the putative Vif-resistant control protein, A3G-4KR, in which the polyUb residues previously described Iwatani *et al.* were changed to arginine, was only partially resistant to HIV Vif (**Figure 4-1B**). Thus, A3G may also contain alternative sites of polyUb.

Multiple Internal Lysine Residues in the A3F N- and C-termini Are Suitable Substrates for Vif-dependent polyUb – To map the individual lysines in A3F available for polyUb,

we reverted each of the 19 R residues of A3F-19KR back to K and assessed the sensitivity of these mutants to Vif in comparison with the parent A3F-19KR in a single cycle of replication. Seven individual lysine reversions in A3F-19KR – R40K, R52K in Region 1; R209K in Region 2; and R334K, R337K, R355K, R358K in Region 3 – induce a statistically significant increase in infectivity in the presence of Vif over the parent lysine-free variant (**Figure 4-2A-C**). In addition to these seven residues, it is possible that others in each region might also be functional targets since we noted some variability from experiment to experiment in which R-to-K changes appeared to sensitize A3F-19KR to Vif, the extent of which is indicated by the associated error bars. For example, A3F-19KR R185K appeared Vif-susceptible in the experiment from which the Western blots shown in **Figure 4-4B** derive but failed to yield a statistically significant increase in infectivity over A3F-19KR when averaged over three, independent experiments. This implies that, despite the preferential targeting of certain residues in A3F, the process of Vif-mediated polyUb may be sufficiently promiscuous to find alternative sites with lower efficiency when preferred options are not available. Such a semi-stochastic character is consistent with the distribution of more consistently targeted residues throughout both the N- and C-termini of A3F as well as the generally more efficient degradation of control mutants in which the lysine residues in all but one linear region have been changed to arginine, yielding six or seven potential targets of polyUb in that region rather than the one found in associated individual revertants. The degradative Vif-mediated polyUb of A3F is therefore not a structurally fixed Ub event such as those associated with regulatory functions or that previously described for A3G [*e.g.* (Batonnet *et al.*, 2004; Das-Bradoo *et al.*, 2010; Iwatani *et al.*, 2009; Scherer *et al.*, 1995; Wu *et al.*, 2003; Zheng

et al., 2002)], but rather may be more accurately described as a flexible sampling of available substrates wherein at least seven residues throughout both the N- and C-termini of A3F can be polyubiquitinated despite binding of Vif at the protein C-terminus (Albin *et al.*, 2010b; Russell *et al.*, 2009b; Smith and Pathak, 2010; Zhang *et al.*, 2008).

Mass Spectrometric Identification of Lysine Acceptor Sites in A3F and A3G – Previous analyses have relied exclusively on genetic and biochemical techniques for the determination of Ub sites in A3G (Dang *et al.*, 2008b; Iwatani *et al.*, 2009; Shao *et al.*, 2010; Wang *et al.*, 2011b). To acquire direct, relatively unbiased biophysical evidence for Ub of the lysine residues identified as functional targets in A3F and to determine whether sites in A3G beyond those previously identified might be modified by Ub, we subjected A3F and A3G to ubiquitin remnant profiling, a technique whereby a monoclonal antibody specific for the K-GG motif characteristic of trypsinized, Ub peptides is used to immunoprecipitate the Ub proteome in a cell lysate prior to mass spectrometric analysis (Xu *et al.*, 2010). Peptides harboring Ub on residues 52, 234, 334, 352, 355 and 358 of A3F were highly enriched in the presence of Vif, indicating substantial modifications at the A3F C-terminus despite the binding of Vif to the same end of the protein (**Figure 4-3A**). Moreover, despite a few minor discrepancies in the specific acceptor sites implicated by the mass spectrometric and genetic datasets, the results of these approaches largely overlap and collectively reinforce the notion that the modification of APOBEC3 proteins is not restricted to a structurally rigid subset of lysine acceptor sites (see **Discussion**).

For A3G, peptides corresponding to ubiquitination of five N-terminal and five C-terminal residues were enriched in the presence of Vif (**Figure 4-3B**). These data again implicate residues of both the N- and C-termini as Ub acceptor sites and are further consistent with the inability of K-to-R mutations at the residues previously identified by Iwatani *et al.* to render A3G fully resistant to Vif (**Figure 4-1B**). Importantly, like the variation observed in the analysis of single K-to-R revertants in A3F-19KR, we found that the exact repertoire of modified sites as determined by mass spectrometry differed slightly from experiment to experiment. For example, a previous analysis of A3G by this same technique identified modification only at residues 297 and 303 among the four residues reported by Iwatani *et al.* while also finding no modification at residue 63 (**Figure 4-3B** and see **Discussion**).

Lysine Residues in A3F and A3G Cluster at Distinct Predicted Surfaces – To determine whether the susceptibility patterns observed in **Figures 4-2** and **4-3** may have a structural basis, we created a model of the A3F N-terminus to complement the C-terminus model previously reported (Albin *et al.*, 2010b). This yielded the surprising finding that in both the N- and C-termini, the lysine residues in A3F typically cluster at one surface, which in the case of the C-terminus is opposite to the surface strongly implicated in the interaction with Vif (**Figure 4-4A-B**). Thus, the placement of lysines within A3F is consistent with the promiscuous ability of the Vif-E3 ligase complex to promote polyUb at multiple residues in A3F since these residues are, broadly speaking, structurally comparable. It is important to note, however, that structural data are not yet available to make a robust model of full-length A3F. When the Ub acceptor sites implicated in **Figure 4-3B** are

visualized on our previously-described full-length model of A3G, however, we observe a similar structural clustering of lysine residues at a surface opposite that bound by Vif (**Figure 4-4C**).

Lack of evidence for polyUb at sites other than internal APOBEC3 lysines – Although our results above as well as those of Iwatani *et al.* provide positive evidence for internal lysines as the predominant sites of polyUb in APOBEC3 proteins, others have previously reported that either Vif itself or the N-terminus of A3G may be important sites of polyUb en route to A3G degradation (Dang *et al.*, 2008b; Shao *et al.*, 2010; Wang *et al.*, 2011b). The results of Dang *et al.* implicating autoUb of Vif as the functionally relevant modification for A3G degradation may reflect the targeting of lysine residues in the C-terminal tag of the otherwise lysine-free A3G variant used [(Dang *et al.*, 2008b; Iwatani *et al.*, 2009) and consistent with sequencing of constructs we received from Dr. Yong-Hui Zheng, Michigan State University]. To functionally address the possibility that A3F or A3G may be N-terminally modified, however, we sought to create variants that would be predicted to be resistant to N-terminal modification.

Changing residue two of A3F and A3G from K to P, G or V favors cleavage of the initiating methionine to create a new N-end [*e.g.* (Varshavsky, 2011)]. Among potential N-terminal residues, the imino group of P is biochemically distinct and likely resistant to N-terminal Ub since, unlike all other amino acids, artificial fusion of Ub to an N-terminal P yields a bond resistant to the action of the deubiquitinating enzymes present in yeast and in mammalian cell lysates [*e.g.* (Bachmair *et al.*, 1986; Gonda *et al.*, 1989)]. The resistance of P to N-terminal acetylation also obviates potential complications

associated with the ability of this modification promote degradation (Hwang *et al.*, 2010), while G and V are subject to modification by N-terminal acetylases distinct from those predicted to act on wildtype A3F and A3G [*e.g.* (Arnesen, 2011)]. We therefore predicted that changing the second APOBEC3 amino acid to P, G or V should alter the Vif sensitivity of the resultant variant if Vif regulates APOBEC3 proteins via modifications of the APOBEC3 N-terminus. We observed, however, that changes to the second amino acid of APOBEC3 proteins do not substantially alter their Vif susceptibility in a single cycle of replication (**Figure 4-5**). Combined with positive evidence for the Ub of internal lysines, we therefore conclude that the APOBEC3 N-terminus is not a major target of Vif function.

Discussion

We present here genetic and biophysical evidence for the polyUb of multiple internal lysines distributed throughout both the N- and C-termini of A3F and A3G (**Figures 4-1, 4-2, 4-3**). The robust flexibility of the Vif-E3 ligase complex to promote the polyUb of APOBEC3 proteins at a wide range of acceptor sites is consistent with an active sampling of APOBEC3 lysines by the associated E2~Ub conjugate. Such a sampling model is further consistent with the observation that single R-to-K reversions in A3F-19KR context are generally less efficient in sensitizing A3F-19KR to Vif relative to control mutants in which all the available substrates of a given region are lysines, particularly in Region 1 (**Figure 4-2A**). This relatively indiscriminate process makes intuitive sense since the default mode for Vif should be to rapidly degrade APOBEC3 proteins with no apparent need for tight regulation of the kind mediated by more selective Ub events [*e.g.* (Batonnet *et al.*, 2004; Das-Bradoo *et al.*, 2010; Iwatani *et al.*, 2009; Scherer *et al.*, 1995; Wu *et al.*, 2003; Zheng *et al.*, 2002)].

Consistent with the promiscuous modification of A3F lysines, we observe that mutation of the residues previously reported by Iwatani *et al.* as the sole functional targets of Vif-mediated polyUb in A3G render the protein only partially Vif-resistant, while many N- and C-terminal lysines in A3G are modified in the presence of Vif (**Figures 4-1A** and **4-3B**). We therefore think it likely that the antiparallel Iwatani *et al.* model previously proposed for APOBEC3 polyUb does not accommodate all of the data (**Figure 4-6A**). It is possible that the model proposed by Iwatani *et al.* may be partially correct since the A3G-4KR mutant is notably less susceptible to Vif than wildtype A3G, and we do note the presence of some susceptible lysine residues in the A3F N-terminus

(**Figures 4-1B** and **4-2A-B**). Nonetheless, the reversion of multiple lysines in the A3F C-terminus can also render A3F-19KR sensitive to Vif, demonstrating extensive flexibility in the ability of RBX2-activated E2~Ub conjugates to reach dispersed sites.

Rather than the rigid antiparallel model of APOBEC3 polyUb proposed by Iwatani *et al.*, it is more likely that APOBEC3 proteins bind Vif in an angled or perpendicular orientation that would expose to activated E2~Ub the entirety of the lysine-rich surfaces predicted to be opposite the Vif binding sites in A3F and A3G (**Figure 4-6B**). It is also possible that multimeric forms of APOBEC3 proteins or of the E3 Ub ligase complex might yield results similar to ours, although we have no evidence to directly support such a model.

We noted that there was some variability in which specific R-to-K reversions in A3F-19KR sensitized the protein to Vif from experiment to experiment (**Figure 4-2**), while linear regional mutants with multiple sidechains available for polyUb were consistently Vif-sensitive (**Figures 4-1** and **4-2**). Similarly, independent, unbiased mass spectrometric analyses of A3G indicated some variation in the specific residues modified. For example, focusing just on the four residues functionally implicated by Iwatani *et al.*, residues 297, 303 and 334 were identified in one experiment (**Figure 4-3B**), while only residues 297 and 303 were identified in a second and neither experiment produced any evidence for the modification of residue 301. We interpret these data to mean that the Vif-mediated polyUb of APOBEC3 proteins is a semi-stochastic affair in which the associated E2~Ub conjugate is able to sample multiple available sidechains for polyUb. As a result, changing the residues most efficiently targeted for modification may simply force the modification of less efficiently targeted sidechains, consistent with data from

other fields indicating that components of the E3 Ub ligase complex may display substantial structural flexibility [*e.g.* (Deshaies and Joazeiro, 2009; Liu and Nussinov, 2009, 2010, 2011)].

Similarly, there were several discrepancies between the specific residues likely modified according to genetic versus biophysical approaches (**Figures 4-2** and **4-3**). Three A3F residues – 40, 209 and 337 – were found to be significantly targeted in the genetic data but were not identified by mass spectrometry, while the opposite was true of residue 234. This may reflect the fact that the wildtype protein used for mass spectrometry simply has more residues available for polyUb and that, consequently, the wildtype protein is preferentially modified at those residues common to the genetic and mass spectrometric datasets. Alternatively, the structures of A3F and A3F-19KR may differ slightly despite their similar restriction activities (**Figures 4-1** and **4-2**). It is also possible that further experiments may find additional residues modified that are not apparent in the data presented here, as we have observed with A3G. Regardless, the overarching message remains clear: many residues in both the N- and C-terminal domains of both A3F and A3G are suitable substrates for polyUb.

In summary, we report here that the Ub acceptor sites in A3F and A3G are distributed throughout the protein, suggesting that the conjugated E2 enzyme(s) recruited by the Vif-E3 ligase complex are more flexible in their substrate targeting potential than previously proposed and that blocking the modification sites themselves will not be a viable strategy for the therapeutic stabilization of APOBEC3 proteins. It will be interesting, as more structural data become available, to find whether the apparent clustering of lysines at surfaces distinct from those at which Vif interacts is a general

feature of the APOBEC3-Vif interaction. It may be that the positively-charged Vif protein achieves efficient degradation of APOBEC3 proteins, in part, by binding to negatively charged regions of susceptible deaminase domains such as the α4 helix and thereby positioning distinct, lysine-rich surfaces for efficient polyUb in the context of the fully assembled E3 Ub ligase complex.

Materials and Methods

Plasmids – Wildtype A3F and A3G coding sequences are identical to those found in GenBank entries NM_145298 and NM_021822 and previously reported [*e.g.* (Albin *et al.*, 2010a)]. APOBEC3 expression constructs were expressed in pcDNA3.1-derived vectors with a C-terminal V5 tag as previously described (Albin *et al.*, 2010a); in all V5 constructs except the Vif interaction mutants A3F QE323-324EK and A3G D128K, however, the inherent tag lysine was mutated to arginine to prevent artifactual tag Ub such as that previously described (Iwatani *et al.*, 2009). K-to-R mutants were made by sequential site-directed mutagenesis and subcloning; A3G-20KR was graciously provided by Dr. Yong-Hui Zheng (Dang *et al.*, 2008b) and cloned into our lysine-free V5 vector. The Vif-deficient HIV$_{IIIB}$ proviral construct and IIIB Vif-HA expression construct were as previously described (Albin *et al.*, 2010a). All constructs were verified by sequencing and restriction digestion.

Cell lines – HEK293T cells were maintained in DMEM medium supplemented with 10% fetal bovine serum and penicillin-streptomycin. CEM-GFP reporter cells were obtained from the AIDS Research and Reference Reagent Program (Gervaix *et al.*, 1997) and maintained in RPMI medium supplemented with 10% fetal bovine serum, penicillin-streptomycin and β-mercaptoethanol.

Single-cycle infectivity protocol – Single-cycle infectivity experiments were carried out as previously described (Albin *et al.*, 2010a). 250,000 HEK293T cells were plated into 2 mL of supplemented DMEM in 6-well plates. The following day, cells were transfected

156

with 200 ng of either a vector control or an A3-V5 expression construct, 100 ng of a Vif-HA expression construct, and 1.6 μg of replication competent, Vif-deficient HIV$_{IIIB}$ provirus using Trans-IT transfection reagent (Mirus Bio). Two days after transfection, transfected cells were harvested and lysed as described below, and virus-containing supernatants were filtered through 0.45 μm filters and used to infect CEM-GFP cells in 96-well plates. Three days post-infection, target CEM-GFP cells were fixed with 4% paraformaldehyde and analyzed by flow cytometry for viral infectivity on a FACSCanto flow cytometer (BD Biosciences).

Western blot analysis – HEK293T cells transfected as described above were washed in PBS and lysed in 250 μL of a lysis buffer composed of 25 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM MgCl$_2$, 50 μM ZnCl$_2$, 10% glycerol and 1% Triton X-100 and supplemented with 50 μM MG132 and complete protease inhibitor (Roche). A 5x sample buffer consisting of 62.5 mM Tris pH 6.8, 20% glycerol, 2% sodium dodecyl sulfate, 5% β-mercaptoethanol, and 0.05% bromophenol blue was added to each cell lysate to a 2x concentration, and the mixture was boiled for 5-10 minutes and subjected to fractionation by 10% SDS-PAGE and transfer to PVDF membranes. Primary antibodies utilized include mouse anti-V5 (Invitrogen), mouse anti-HA.11 (Covance), and mouse anti-tubulin (Covance). Between probing with different primary antibodies, membranes were stripped with 62.5 mM Tris pH 6.8, 2% SDS, and 100 mM β-mercaptoethanol at 50°C and washed in PBS 0.01% Tween.

Mass spectrometry – HEK293T cells were labeled using a SILAC approach (Ong *et al.*, 2002). Cells were cultured in either "light" SILAC media containing the normal complement of amino aids or "heavy" SILAC media wherein lysine was replaced with $^{13}C_6$-lysine. FBS was dialyzed in both light and heavy media formulations to remove free unlabeled amino acids. Four 15 cm plates of light media each were transfected with 2.5 µg Vif-Strep and 2.5 µg of either A3F-V5 or A3G-FLAG. An additional four 15 cm plates of 293T cells cultured in heavy media each were also transfected with 2.5 µg pcDNA4 vector and 2.5 µg either A3F-V5 or A3G-FLAG. Cells were detached from the plate with 10 mM EDTA in PBS, spun down, and snap frozen. The frozen cell pellets were lysed in buffer containing 8M urea, 0.1M Tris pH 8.0, 150 mM NaCl, and a combination of protease inhibitors (Complete tablet, Roche). Protein concentration was measured by Bradford assay (QuickStart 1x Reagent, Bio-Rad). 50 mg protein lysate from equal portions of light and heavy culture conditions for A3F-V5 or A3G-FLAG, respectively, were combined and then subjected to reduction with 4 mM TCEP for 30 minutes at room temperature, alkylation with 10 mM iodoacetamide for 30 minutes and room temperature in the dark, and overnight digestion with 250 µg of trypsin (Promega) at room temperature. Treated lysates were subsequently desalted using SepPak C18 cartridges (Waters), lyophilized for two days and immunoprecipitated using an antibody specific for the K-GG motif characteristic of trypsinized, Ub-modified peptides (UbiScan, Cell Signaling Technology). Immunoprecipitates were then desalted using C18 STAGE tips (Thermo Scientific), evaporated, and analyzed in duplicate with a two-hour gradient on an Orbitrap Elite Mass Spectrometer (Thermo Scientific). The reverse phase gradient was delivered by an Easy nLC 1000 liquid chromatography system (Thermo

Scientific) from 5% to 30% acetonitrile in 0.1% formic acid. Columns used were a 100 µm x 2 cm pre-column packed with 5 µm ReproSil Pur C18 particles and a 75 µm x 10 cm analytical column packed with 2 µm ReproSil Pur C18 particles (Thermo Scientific). The Orbitrap Elite continuously collected data in data-dependent acquisition mode, acquiring a full scan in the orbitrap at 120,000 resolution followed by collision-induced fragmentation of the top 20 most intense peaks from full scan in the ion trap. Dynamic exclusion was enabled to exclude repeated fragmentation of peaks for 30 seconds. Charge state screening was enabled to reject fragmentation of unassigned or singly charged species. Results were analyzed using the MaxQuant software package (Cox and Mann, 2008).

Homology modeling of the A3F N- and C-terminal deaminase domains – Modeling of the A3F CTD has been described previously (Albin *et al.*, 2010b). The A3F NTD model including residues 1–190 was generated using YASARA (Krieger *et al.*, 2002) with the crystal structure of A3G191–384 2K3A (PDB 3IR2) as a template (Shandilya *et al.*, 2010). Alignment with the template sequence was iteratively optimized using SwissProt and TrEMBL sequences, the predicted secondary structure and the three dimensional structure of the template. Refinement of the model was carried out using knowledge-based and electrostatic interactions in unrestrained molecular dynamics with explicit solvent molecules. Insertions were accounted for by searching the Protein Data Bank for superimposable loop ends with model anchor points and optimized for energy minimization.
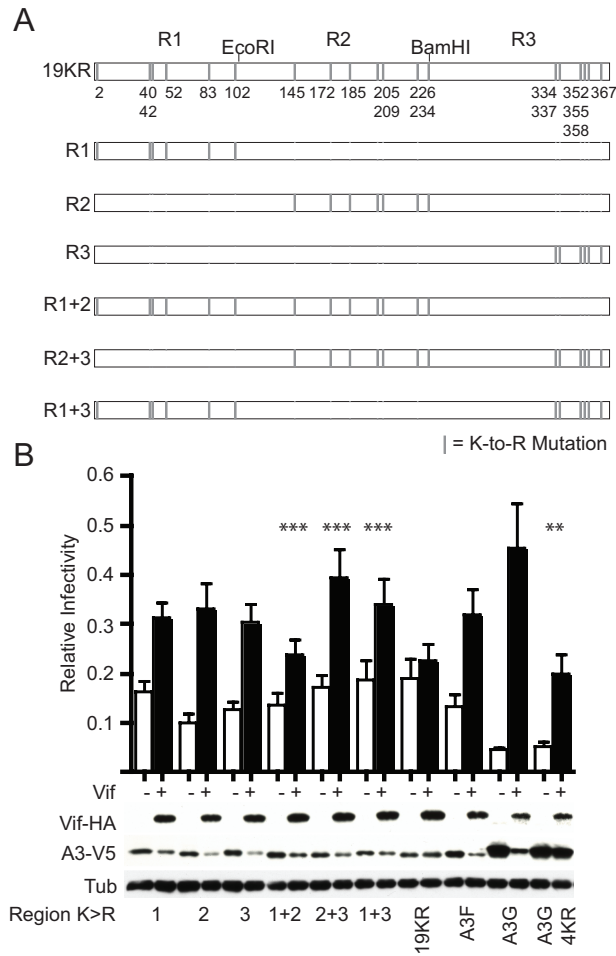
Figure 4-1. The sites of functional Vif-mediated polyUb in A3F are distributed throughout the protein.

Figure 4-1. The sites of functional Vif-mediated polyUb in A3F are distributed throughout the protein. (A) A schematic showing the K-to-R variants of A3F tested in **Figure 4-1B**. **(B)** Single-cycle infectivity of Vif-deficient viruses produced in the presence of the indicated A3F-V5 variant +/- transcomplementation by Vif-HA. Infectivity data represent the mean and SEM of four independent experiments; here as in all singe-cycle infectivity experiments conducted, infectivity is normalized to the infectivities of virus produced in the presence of an APOBEC3 vector control -/+ Vif. Statistics were derived by carrying out one-tailed, paired T tests comparing the infectivity of virus produced in the presence of each APOBEC3 construct -/+ Vif. ** = $p < 0.05$; *** = $p < 0.01$. Western blots showing the steady state levels of Vif-HA and A3-V5 are derived from the producer cells from one of the experiments conducted.
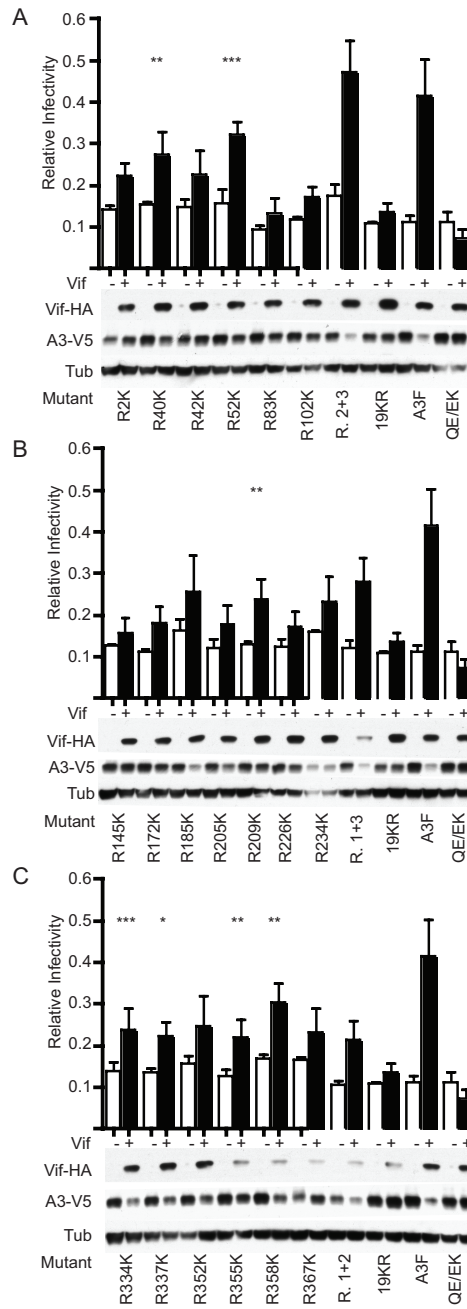
Figure 4-2. Multiple internal lysine residues in A3F are suitable substrates for functional Vif-dependent polyUb.

Figure 4-2. Multiple internal lysine residues in A3F are suitable substrates for functional Vif-dependent polyUb. (A-C) Single-cycle infectivity of individual single lysine revertants in an A3F-19KR background in the absence or presence of Vif where data represent the mean and SEM of three independent experiments with associated Western blots showing producer cell steady state levels of Vif-HA and A3-V5 derived from one of these experiments. Statistics represent two-way ANOVA with Dunnett's post-test comparing the infectivity recovery of each single amino acid revertant (Vif+/Vif-) to that of the A3F-19KR control. * = $p < 0.1$; ** = $p < 0.05$; *** = $p < 0.01$. Controls that appear in multiple panels of this figure (*e.g.* A3F in **A**, **B** and **C**) are regraphed and reblotted for visual comparison with the mutants of each individual region; all data shown in this figure are the result of one large experiment repeated three times and presented visually as three parts.

A

| Residue: | 52 | | 234 | *289 | *324 | 334 355 352 | 358 |
| Enrichment: | 21+ | | ++ | | 40 | 43+ 32 33+ | |

B

| Residue: | 42 52 63 | *128-130 150 163 | 249 270 297 | 303 334 |
| Enrichment: | 4 5 5 | 8 8 | 7 7 24 9 | ++ |
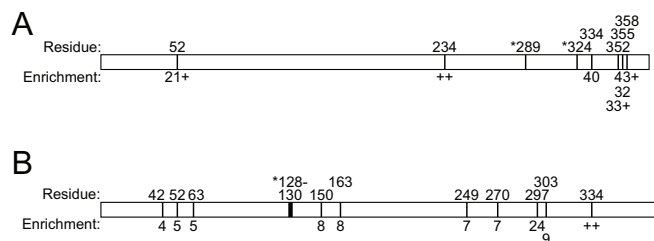
Figure 4-3. Mass spectrometry indicates Ub of multiple internal lysine residues in the N- and C-termini of both A3F and A3G. (A-B) Diagrams showing Ub-modified residues in A3F **(A)** or A3G **(B)** and their associated fold enrichment in the presence of Vif. + Indicates that some of the peptides detected indicating modification of the associated residues were so rare in the absence of Vif as to preclude calculation of a fold-enrichment; in this case, the number represents the fold enrichment from those peptides for which is was calculable. ++ is used for residues where all of the peptides indicating modification of a particular residue precluded calculation of a fold-enrichment.

164
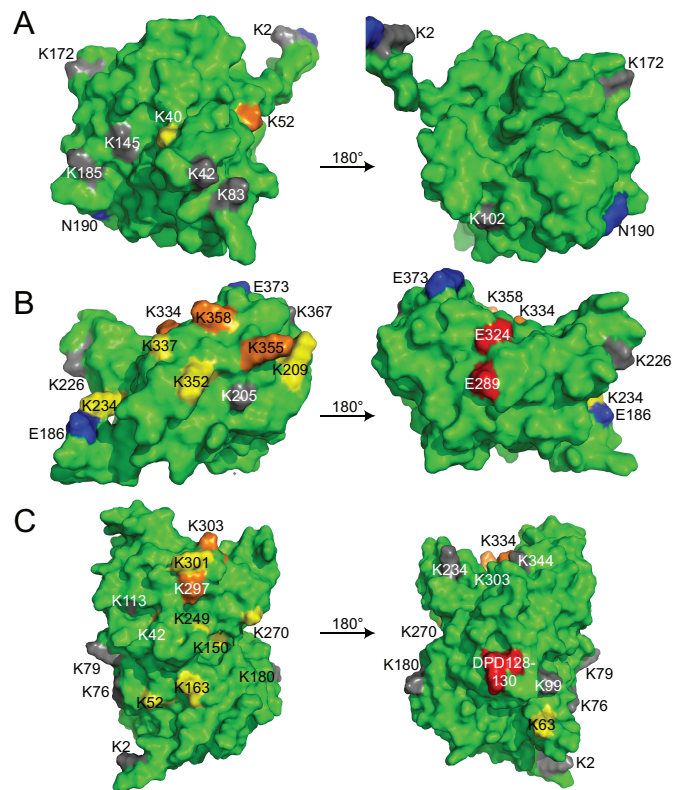
Figure 4-4. Lysine residues in the N- and C-termini of A3F and A3G cluster at distinct predicted surfaces. (A) A model of the A3F N-terminal deaminase domain rotated 180° about the y-axis. **(B)** A model of the A3F C-terminal deaminase domain rotated 180° about the y-axis. **(C)** A previously-described model of full-length A3G rotated 180° about the y-axis (Harjes *et al.*, 2009). Orange indicates residues implicated by **Figures 4-2** and **4-3**, while yellow indicates residues implicated by one or the other; in **(C)**, A3G residues implicated by Iwatani *et al.* and also identified in **Figure 4-3B** are orange, while additional residues not implicated by Iwatani *et al.* are yellow. Lysine residues not significantly implicated by either data set are gray; Vif interaction residues are red (Albin *et al.*, 2010; Albin *et al.*, 2010b; Smith and Pathak, 2010).
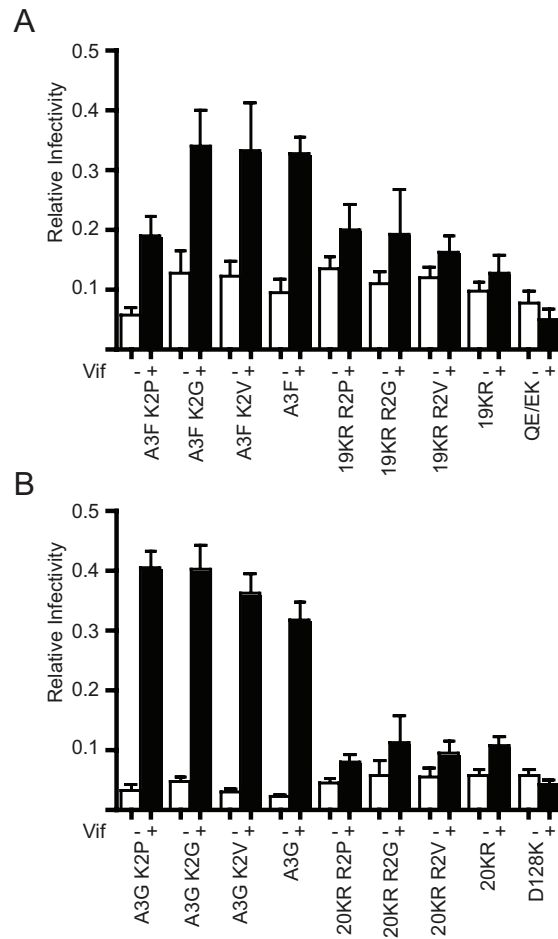
Figure 4-5. Changes at the APOBEC3 N-terminus do not alter Vif susceptibility. (A-B) Single-cycle infectivity experiments demonstrate the Vif sensitivity of A3F, A3G and lysine-free variants thereof to Vif when the second amino acid is altered as shown. Infectivity data represent the mean and SEM of four independent experiments.
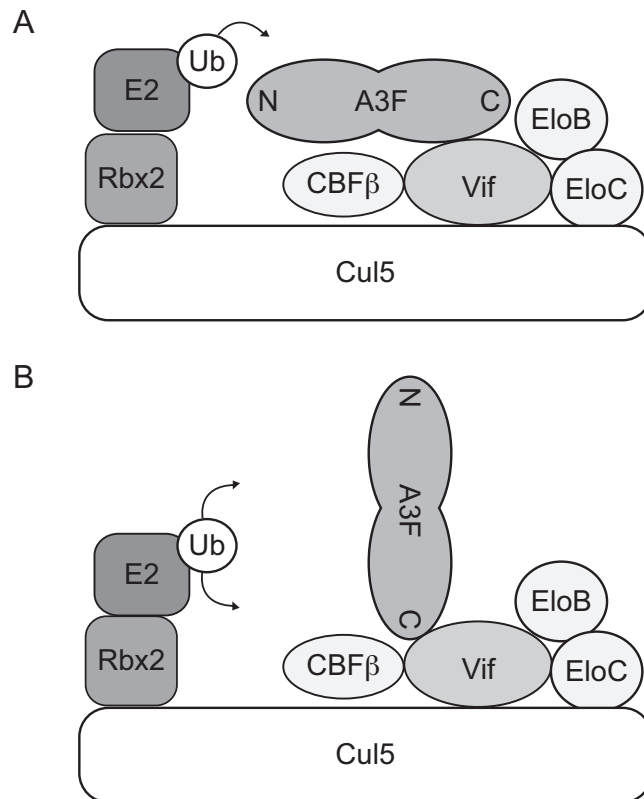
Figure 4-6. An alternative model of APOBEC3 binding to the Vif-E3 ligase complex.

(A) The antiparallel model previously proposed by Iwatani *et al.* in which binding of Vif at one deaminase domain orients an APOBEC3 protein for polyUb at its second deaminase domain (Iwatani *et al.*, 2009). **(B)** An alternative, angled/perpendicular model for A3F binding to the Vif-E3 ligase complex in which binding of Vif to a lysine-poor surface orients the lysine-rich surfaces of A3F for polyUb.

Postscript

This manuscript is presently *in revision*. While we report specific ubiquitin acceptor sites that are more likely targeted than others in A3F and A3G, the broader point suggested by our data is that there is little structural rigidity associated with the ability of an E2-Ub conjugate to find a suitable substrate in E3 ubiquitin ligase complex-associated A3F. It will therefore be of interest to those in the field to know that the model proposed by Iwatani *et al.*, despite its appeal, does not hold for other APOBEC3 proteins – or perhaps even for A3G itself. Our data may further be of use in the functional evaluation of any future co-crystal structures of the full APOBEC3-Vif-E3 ubiquitin ligase complex since the clustering of lysines at surfaces distinct from those at which Vif binds implies a distinctive topology that we anticipate with be borne out structurally.

Chapter 5: Evidence for a Conserved Structural Feature in HIV Vif that Interacts with APOBEC3 Proteins

Foreword

This chapter is a working draft of the following manuscript *in preparation*:

Albin JS, Refsland EW, Ikeda T, Holten JR, Anderson JS and Harris RS. "Evidence for a Conserved Structural Feature of HIV Vif that Interacts with APOBEC3 Proteins." *In preparation*

The data herein are an extension of **Chapter 3** in which we seek to identify the specific amino acids in Vif that interact with the known single amino acid determinants of Vif susceptibility found in the α4 helix of human APOBEC3 proteins. While we are still pursuing confirmatory evidence, we believe the data in this chapter support a novel conception of the APOBEC3-Vif interaction. That is, the residues that we have identified suggest that the interaction of HIV Vif with human APOBEC3 proteins, rather than being three different interactions depending on the substrate APOBEC3 as currently conceived, is defined by a common core contact with a region of Vif containing a conserved GxxxG helix-helix interaction motif, a concept we have dubbed the Trinity Hypothesis.

Summary

The human immunodeficiency virus virion infectivity factor (Vif) links APOBEC3 cytosine deaminases to an E3 ubiquitin ligase complex to mediate their proteasomal degradation and thereby protect viral genomic integrity from the mutations created by these restriction factors. The APOBEC3-Vif interaction is thus an attractive potential target for the development of small molecule therapeutics that might liberate APOBEC3 proteins to inhibit HIV. Our knowledge of the molecular details of the interactions of APOBEC3 proteins with Vif, however, remains limited. Here, we report the results of genetic selection experiments in which we have evolved compensatory changes in Vif that counteract the Vif-resistant APOBEC3 variations of the α4 helix. Together, these data identify a putative α-helix in HIV Vif containing a GxxxG helix-helix interaction motif and three conserved glycine residues in a stretch of 11 amino acids that mediate compensatory changes in Vif to counteract Vif-resistant APOBEC3 proteins. Thus, we propose an antiparallel helix-helix interaction between the APOBEC3 α4 helix and a putative glycine-rich helix in HIV Vif as the common structural core of APOBEC3-Vif interactions.

Introduction

When APOBEC3 proteins are encapsidated into budding HIV virions in producer cells, they can inhibit HIV reverse transcription in target cells by introducing massive levels of C-to-U transitions into minus-strand viral cDNA during reverse transcription. [*e.g.* (Harris *et al.*, 2003a; Mangeat *et al.*, 2003; Zhang *et al.*, 2003)]. The net effect of this is decreased viral cDNA accumulation and the integration of hypermutated, and presumably hypofunctional, proviruses. To prevent encapsidation and thereby prevent the deleterious effect of APOBEC3 proteins on the viral genome, the HIV accessory protein virion infectivity factor (Vif) links APOBEC3 proteins to an E3 ubiquitin ligase complex, facilitating the polyubiquitination and proteasomal degradation of these restriction factors (Conticello *et al.*, 2003; Marin *et al.*, 2003; Mehle *et al.*, 2004b; Sheehy *et al.*, 2003; Stopak *et al.*, 2003; Yu *et al.*, 2003).

While the APOBEC3-Vif interaction has garnered substantial interest due to its potential as a novel antiretroviral therapeutic target, our understanding of the interactions between APOBEC3 proteins and Vif remains limited. Many structures of the A3G C-terminus have been published, and although these are useful for modeling other APOBEC3 proteins, the A3G C-terminus itself is not targeted by Vif (Chen *et al.*, 2008; Furukawa *et al.*, 2009; Harjes *et al.*, 2009; Holden *et al.*, 2008; Li *et al.*, 2012; Shandilya *et al.*, 2010). More recently, a crystal structure of the Vif-susceptible APOBEC3 protein A3C has been reported along with extensive mutagenesis identifying novel components of the Vif recognition region applicable to A3C, A3D and A3F, and our own collaborative efforts have contributed to a crystal structure of the A3F C-terminus [(Kitamura *et al.*, 2012; Bohn *et al.*, *in preparation*) and see **Supplementary Chapter 1**].

Genetically, several single amino acid determinants that can render APOBEC3 proteins Vif-resistant have been identified by comparative mutagenesis of innately Vif-resistant and Vif-susceptible cytosine deaminase domains, including but not limited to D128 and D130 in A3G (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Lavens *et al.*, 2010; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004), D121 in A3H (Zhen *et al.*) and E289 and E324 in A3F [(Albin *et al.*, 2010b; Russell *et al.*, 2009b; Smith and Pathak) and see **Chapter 3**]. These studies have generally been interpreted as implying the existence of distinct Vif interaction regions in different APOBEC3 proteins since, in the case of A3G and A3F, the Vif interaction residues are linearly distinct, being found in the N- or C-terminal deaminase domains of these double domain APOBEC3 proteins, respectively (Russell *et al.*, 2009b; Smith and Pathak, 2010). If one considers the APOBEC3 proteins phylogenetically, however, a pattern emerges in which all of these critical determinants structurally cluster on or adjacent to the α4 helix of a Vif-susceptible cytosine deaminase domain (Albin *et al.*, 2010b). Thus, it is possible that Vif recognizes a common structural determinant in APOBEC3 proteins centered on this conserved helix.

Progress on the Vif side of this interaction has been hampered by the difficulty of purifying quantities of Vif amenable to structural study. To date, published structural assessments of Vif are limited to a short Vif peptide of the conserved BC box co-crystallized with Elongin C (Stanley *et al.*, 2008), NMR characterization of Vif assembly with the E3 ubiquitin ligase complex (Bergeron *et al.*, 2010) and mass spectrometric characterization of crosslinked multimeric forms of Vif (Auclair *et al.*, 2007). It is possible that the recent discovery of Vif cofactor CBFβ may soon improve this situation

given the potential to purify greater quantities of the notoriously difficult viral accessory protein by coexpression (Jäger *et al.*, 2011; Zhang *et al.*, 2011), but for the moment, much work remains to be done toward understanding the structural features that translate into Vif function.

In contrast, a number of genetic studies on the functionality of various Vif mutants against A3F, A3G or both has yielded, in recent years, an abundance of data summarized in **Table 1-2** (Albin *et al.*, 2010a; Chen *et al.*, 2009; Dang *et al.*, 2010a; Dang *et al.*, 2010b; Dang *et al.*, 2009; Fourati *et al.*, 2010; He *et al.*, 2008; Mulder *et al.*, 2010; Nagao *et al.*, 2010; Pery *et al.*, 2009; Russell and Pathak, 2007; Simon *et al.*, 1999; Simon *et al.*, 2005; Tian *et al.*, 2006; Yamashita *et al.*, 2008; Yamashita *et al.*, 2010; Zhang *et al.*, 2008). An attempt to illustrate some of these data in broad strokes is provided in **Figure 1-4**, wherein one can identify certain regions such as amino acids 14-DRMR-17 and 74-TGERDW-79 as critical for interaction with A3F, others such as K26 and 40-YRHHY-44 as critical for interaction with A3G and others such as amino acids in the range of 55-72 as critical for interaction with both APOBEC3 proteins. Taken as a whole, this implies that the APOBEC3-Vif interaction is structurally complex since disparate parts of Vif are associated with the ability to neutralize one or more APOBEC3 proteins.

Even such attempts at distillation are perhaps less informative than one might imagine, however, because all of the aforementioned studies identified only loss of function mutants that could be the result of any number of nonspecific changes to Vif structure rather than an actual, direct interaction. The phenotypic data themselves are at times contradictory (**Table 1-2**), and there are instances in which even well-established

features of current conceptions of the APOBEC3-Vif interaction break down. For example, *vif* genes with mutations of amino acids 14-DRMR-17 are generally accepted as selectively susceptible to A3F but not A3G [*e.g.* (Russell and Pathak, 2007)], yet alteration of this same region to 14-SEMQ-17 on the basis of comparative sequence analysis between HIV and SIV Vifs permits the functional neutralization of the HIV Vif-resistant proteins rhA3G and human A3G D128K by an otherwise unaltered HIV Vif (Schröfelbauer *et al.*, 2006). Mutations in other regions such as 40-YRHHY-44, which selectively displays strong loss of function versus A3G under conditions of transient overexpression may nevertheless fail to yield similar phenotypes in a more physiologic spreading infection system [(Russell *et al.*, 2009a) and **Supplementary Chapter 2**, but contrast with (Yamashita *et al.*, 2008)], raising the question of whether such regions are as critical as they initially appeared (Russell and Pathak, 2007).

Despite the diversity of Vif interactions with APOBEC3 proteins implied by **Table 1-2**, the placement of most of the known single amino acid determinants of Vif susceptibility within a single, defined structural feature of APOBEC3 proteins strongly implies that the seemingly distinct interactions of Vif with APOBEC3 proteins must have some feature in common. Moreover, this common feature must depend on interaction with negatively charged residues of the α4 helix since all of the known single amino acid determinants of Vif susceptibility in this region can render an APOBEC3 protein Vif-resistant via a negative-to-positive charge change. If a specific residue in Vif directly interacts with one of these critical amino acid determinants, then, one would predict that viruses adapted to growth in the presence of these Vif-resistant APOBEC3 proteins should acquire compensatory net-negative charge changes that permit functional

interaction with the newly positively charged interaction node in a given APOBEC3 protein, a concept illustrated in **Figure 5-1A**.

To define such directly interacting residues, we have adapted Vif-proficient HIV molecular clone-derived viruses of defined sequence to growth in the presence of the Vif-resistant APOBEC3 proteins A3F QE323-324EK, rhesus macaque A3F (rhA3F), A3G DPD128-130KPK or rhA3G and determined the *vif* genotypes of the resultant adapted virus cohorts. Surprisingly, the *vif* changes accumulated through this process displayed a strong propensity to scatter throughout the gene, with some isolates displaying no discernible changes in *vif*, potentially implying the existence of alternative, Vif-independent mechanisms of APOBEC3 neutralization available to HIV (see **Chapter 6**). Within *vif*, however, we identified compensatory changes capable of preserving viral infectivity in the presence of APOBEC3 proteins resistant to the degradative effects of the parental Vif, namely changes G71D and G82D, which allow HIV Vif to overcome the Vif-resistance mutations A3F E324K and A3G D130K, respectively.

On comparing sequences, we noted that the phylogenetic equivalents of these residues in APOBEC3 proteins are separated by the same 11 amino acid distance found between the corresponding compensatory mutations in Vif. Moreover, this 11 amino acid stretch of Vif carries in it a 71-GxxxG-75 helix-helix interaction motif, which when modeled as a helix places four, conserved, glycine-negatively charged residue pairs at 71, 75, 78 and 82 on the same surface of Vif available for interaction with the APOBEC3 α4 helix.

Taken together, our data implicate a conserved, antiparallel helix-helix contact between the APOBEC3 α4 helix and a putative glycine-rich Vif helix as a shared feature

of APOBEC3-Vif interactions. These data represent the first identification of directly interacting residues involved in the APOBEC3-Vif interaction and provide insight into the structural basis of this interaction that has to date eluded structural study. Perhaps more importantly, our data suggest that the extensive array of interactions described by **Table 1-2** can be distilled into a single target that may be highly susceptible to therapeutic intervention.

Results

HIV can adapt readily to growth in the presence of Vif-resistant APOBEC3 proteins –
The Vif proteins produced by the lentiviruses infecting a given species are thought to be
specially adapted to neutralize the APOBEC3 repertoire of that species (Bogerd *et al.*,
2004; LaRue, 2010; Mangeat *et al.*, 2004; Mariani *et al.*, 2003; Schröfelbauer *et al.*,
2004; Xu *et al.*, 2004). As a consequence, lentiviral replication is typically restricted by
APOBEC3 proteins from non-cognate species since Vif has not evolved for activity
against the novel host APOBEC3 repertoire. We have previously validated this
phenomenon using rhesus macaque A3F and A3G as well as derivatives of human A3F
and A3G that have been engineered to resist Vif through the incorporation of specific
rhesus macaque residues into the Vif interaction site of the human protein, QE323-
324EK in A3F and DPD128-130KPK in A3G [(Albin, 2010; Hultquist *et al.*, 2011) and
see **Materials and Methods**]. In principle, then, infection of these cell lines with the
human APOBEC3-adapted HIV Vif acts as an inverse mimic of the APOBEC3-Vif-
dependent components of a cross-species transfer event analogous to those transmissions
from primate to human that have given rise to the HIV epidemic.

Given time to evolve sequence diversity, however, the transfer of a given virus
such as HIV to a novel host species should be possible despite APOBEC3-dependent
barriers since the novel host APOBEC3 repertoire will select variants from among
transmitted genotypes that effectively neutralize the novel APOBEC3 repertoire.
Moreover, the compensatory changes acquired in *vif* should theoretically correspond to
alterations in the APOBEC3 interaction site in Vif that corresponds to the Vif interaction

sites in APOBEC3 proteins that have been either engineered to resist HIV Vif or that are natively HIV Vif-resistant.

To model this situation, we elected to carry out graded selection experiments that would allow sequence diversity to accumulate from defined molecular clone backgrounds while actively selecting for compensatory changes by gradually increasing exposure to any of four Vif-resistant APOBEC3 proteins. APOBEC3 expression levels in the cell lines used are shown in **Figure 5-1B-C**, and the overall strategy is depicted schematically in **Figures 5-1E-F**. In brief, we infected permissive SupT11 derivatives stably transfected with a vector control with viruses derived from molecular clones of defined sequence, HIV$_{\text{IIIB}}$ and HIV$_{\text{LAI-GFP}}$, and gradually passaged these viruses through a series of co-cultures between the SupT11-vector cells and one of four novel species APOBEC3 conditions – A3F QE323-324EK, rhA3F, A3G DPD128-130KPK and rhA3G – before ending each round on a culture purely of cells stably transfected with a Vif-resistant APOBEC3 protein. The exact combinations of virus and cell line are depicted schematically in **Figure 5-1E**, where each prong represents a 24-well plate in which 22 wells were used for selection and two wells were used for positive and negative controls – virus maintained on the original permissive cell line and an uninfected well of co-cultured cells, respectively – and the cycles described are depicted schematically in **Figure 5-1F**.

In total, we completed three rounds of selection. Isolates were saved from 100% nonpermissive cultures at the end of each round based on indicators of improved replication such as enhanced CEM-GFP infectivity, increased syncytia formation or increased cell death as described in **Materials and Methods**. These viruses were then

subjected to a total of three subsequent, MOI-controlled confirmatory passages on fresh Vif-resistant nonpermissive cells of the kind on which they were originally selected; all viruses saved for subsequent passages were saved exclusively from cell lines of the type on which they were originally selected. The total number of adapted viruses as confirmed through the fourth and final passage of saved isolates present at the end of each round of selection as well as fourth-passage spreading infection curves of representative viruses adapted to the indicated condition are shown in **Figure 5-2A-D**. As expected, the number of adapted isolates gradually increased over time for each selection condition, suggesting that our system is capable of generating both sequence diversity and selective pressure for the maintenance of adapted genotypes.

Cross-resistance patterns among viruses adapted to Vif-resistant APOBEC3 proteins show limited specificity – In addition to confirming the adapted phenotype of selected viruses, the fourth and final passage on cells nonpermissive to the parental viruses from which our isolates derive was utilized for two additional purposes – to phenotype adapted isolates when exposed for the first time to novel Vif-resistant APOBEC3 proteins to which they had not been previously adapted and to save infected genomic DNA for sequence characterization of the *vif* alleles present in each adapted virus population. A nearly comprehensive rendering of these sequence and cross-resistance characteristics for each isolate derived from the A3F QE323-324EK and A3G DPD128-130KPK as well as the rhA3F and rhA3G conditions is provided in **Tables 5-1** through **5-8**.

Phenotypically, most isolates from the fourth passage of viruses adapted to human A3F QE323-324EK or to rhA3F displayed strong cross-resistance to each other as shown

for representative isolates in **Figures 5-3A-B** and **5-3I-J** and indicated in **Tables 5-1** and **5-2** and **5-5** and **5-6**. These viruses also often displayed improved growth in the presence of human A3G DPD128-130KPK and rhA3G, although these phenotypes were generally weaker on the former cell lines, often yielding low peaks at late time points in the experiment (**Tables 5-1-2** and **5-5-6** and **Figure 5-3C-D**).

A true appreciation for the extent of such cross-resistance was not feasible for the A3G DPD128-130KPK and rhA3G combinations because rhA3G is relatively poorly restrictive to Vif-proficient HIV; unaltered HIV$_{IIIB}$ viruses and, to a lesser extent, HIV$_{LAI-GFP}$ will usually peak on rhA3G-expressing cells by approximately 2-3 weeks post-infection when starting from low MOI (data not shown). Generally speaking, however, viruses adapted to growth in the presence of A3G DPD128-130KPK were quite proficient at growth in the presence of rhA3G, but the reverse was not true (**Tables 5-3** and **5-4** and **5-7** and **5-8** and **Figures 5-3G-H** and **5-3K-L**). This is likely due to the poorly restrictive phenotype displayed by rhA3G; in this case, it is possible that mundane fitness enhancement or other nonspecific changes in the virus could enable enhanced replication in the presence of rhA3G without readily translating into growth in the presence of A3G DPD128-130KPK, which is strongly restrictive in our spreading infection system.

Surprisingly, when we exposed viruses adapted to A3G DPD128-130KPK to the strongly restrictive A3F QE323-324EK and rhA3F cell lines, these viruses often displayed robust growth phenotypes similar to those observed on their parent cell lines, indicating a lack of specificity in the adaptive features acquired by these viruses through sequential passage relative to viruses similarly passaged in A3F QE323-324EK- or rhA3F-expressing cells (**Tables 5-3** and **5-4** and **5-7** and **5-8** and **Figures 5-3E-F**).

180

A3F QE323-324EK and rhA3F show strong selection for a common mutation despite prima facie *evidence that vif mutations do not explain the adapted phenotypes of all viruses* – To better characterize the *vif* genes present in the viruses adapted to our various Vif-resistant APOBEC3 conditions, we isolated genomic DNA from at or near the peaks of the fourth confirmatory passage from 1-2 cell lines for each virus and amplified the *vif-vpr* region of each virus as previously described (Albin *et al.*, 2010a). We then purified and sequenced these PCR products. These data are presented in several formats to facilitate digestion and allow full access to the selection dataset for cross-referencing as desired. **Tables 1-1** through **1-8** list the *vif* genotype of each adapted isolate next to its phenotypic scoring, while **Figures 5-4** through **5-11** show alignments for the *vif* genes derived from each combination of starting virus template (HIV$_{IIIB}$ or HIV$_{LAI-GFP}$) and selective condition (A3F QE323-324EK, A3G DPD128-130KPK, rhA3F or rhA3G).

A global overview of the mutations selected by each of our conditions is presented in **Figures 5-12** and **5-13**, where we have plotted the total number of isolates from a given selective condition carrying a particular mutation on the y-axis against each position in the Vif coding sequence on the x-axis. In viewing the data this way, one observes that, while several mutational hotspots occur when viruses are adapted to A3F QE323-324EK, mostly notably G71D and, to a lesser extent, E/D117K/N in isolates derived from both HIV$_{IIIB}$ and HIV$_{LAI-GFP}$ templates and E45D and G185R derived from HIV$_{IIIB}$ or HIV$_{LAI-GFP}$ templates, respectively, there are few patterns to the mutations selected by A3G DPD128-130KPK (**Figure 5-12A-B**). The potential exceptions of interest are the most common change in three isolates derived from both IIIB and LAI-

GFP templates, G82D, as well as changes at residue 15 in the 14-DRMR-17 region that has been previously implicated in allowing HIV Vif to overcome rhA3G and A3G D128K (Schröfelbauer *et al.*, 2006). Similarly, among our rhesus selection conditions, we see a strong selection for the G71D change, particularly among HIV$_{IIIB}$-derived isolates, and essentially no pattern at all among viruses adapted to rhA3G, although once again there is one virus containing a change at residue 15 in the 14-DRMR-17 region **(Figure 5-13A-B)**.

Thus, while our selection protocol produced many changes in *vif*, perhaps the most surprising result in these data is the relative lack of consensus mutations in *vif* associated with each selective combination with the exception of G71D with A3F QE323-324EK and rhA3F. In fact, a total of 11 HIV$_{LAI-GFP}$-derived viruses adapted to A3F QE323-324EK, four HIV$_{IIIB}$-derived viruses adapted to A3G DPD128-130KPK, six HIV$_{LAI-GFP}$-derived viruses adapted to A3G DPD128-130KPK, 10 HIV$_{LAI-GFP}$-derived viruses adapted to rhA3F, three HIV$_{IIIB}$-derived viruses adapted to rhA3G and 11 HIV$_{LAI-GFP}$-derived viruses adapted to rhA3G carried no clear mutations in *vif* (**Tables 5-1-8**). The latter two of these may conceivably be explained by the relative dearth of selective pressure imposed by rhA3G, which is generally only weakly restrictive to HIV Vif-proficient viruses (see **Figure 5-3** above). That said, most of these isolates with the exception of several inactivating mutations found among the HIV$_{IIIB}$-derived viruses adapted to A3G DPD128-130KPK did maintain intact *vif* open reading frames. This strongly implies that HIV can acquire accessory mutations outside of *vif* that can synergize with Vif or even act independently, as it would seem for the inactivating mutations noted, to counteract APOBEC3 proteins in the absence of Vif. While we have

noted a number of *vpr* truncations in the viruses selected, consistent with our prior reports selecting adaptive Vif-independent changes, the cell lines used in the studies described in **Chapter 5** and **Chapter 6** were chosen specifically because their levels of APOBEC3 expression restrict the growth of such *vpr*-deficient variants [(Albin *et al.*, 2010a; Haché *et al.*, 2008) and data not shown)]. Anticipating that other Vif-independent mutations might allow HIV to adapt to the presence of APOBEC3 proteins, however, we carried out similar selection experiments with Vif-deficient viruses in parallel with the studies described here. The results of these studies, described in **Chapter 6**, implicate HIV Env in phenocopying the anti-APOBEC3 effects of Vif, demonstrating unambiguously that Vif-independent adaptation to APOBEC3 proteins is possible.

Selected adapted vif alleles neutralize Vif-resistant APOBEC3 proteins – In considering which mutations to further analyze, the obvious choices for A3F QE323-324EK and rhA3F derivatives were G71D, E45D, E/D117K/N and G185R. Among other prominent peaks, R92K was excluded since it typically occurs as a minor allele (**Tables 5-1** and **5-2**). M189I was also excluded because both this change and G185R also affect the coding sequence of the overlapping *vpr* open reading frame, with R12K being caused by Vif G185R and E17K being caused by Vif M189I. It is possible that these may confound analysis in a proviral context, but it is not clear to what extent the *vpr* open reading frame remains intact in our passaged viruses since loss of *vpr* is a common consequence of long-term passage in culture (Nakaya *et al.*, 1994) and, anecdotally, there are many obvious premature stop codons in *vpr* picked up by our sequencing reads aimed at *vif* (data not shown). Thus, these mutations may represent G-to-A noise in a neutrally

evolving open reading frame. Nevertheless, we were intrigued by the emerging theme of glycine changes among our adapted viruses and so opted to retain G185R for further analysis.

Surprisingly, there were almost no major peaks of sequence evolution evident among viruses adapted to A3G DPD128-130KPK (**Figure 5-3B**). This relative lack of selection for changes in *vif* may indicate that the changes to the Vif interface caused by amino acids DPD128-130KPK are gross alterations. Alternatively, the interaction of Vif with these residues may be partially indirect. Finally, it may simply be that having two mutations independently capable of rendering A3G Vif-resistant created a level of stringency too great for our selection system to handle, as this might require two mutations in close proximity despite the proposed cumulative effects of changes at D130 and D128 in rendering A3G Vif-resistant (Huthoff and Malim, 2007).

Despite the relative lack of activity in *vif* genes adapted to Vif-resistant variants of A3G, we were interested in two changes found among the these viruses. Two viruses contained changes at R15, which was of interest both because these changes were also evident in preliminary clonal sequencing of viruses adapted to rhA3G (data not shown) and because R15 falls within the 14-DRMR-17 region that has been previously mutated to overcome Vif-resistant variants of A3G (Schröfelbauer *et al.*, 2006). More importantly, though, the sole peak among A3G DPD128-130KPK-adapted viruses was G82D. Like G71D, this is a charge change consistent with neutralization of an aberrant positive charge in Vif-resistant A3G variants. Moreover, we noticed that the 11 amino acids separating G71 from G82 in Vif is equivalent to the 11 amino acids separating the phylogenetic equivalents of A3F E324 and A3G D130 [A3F D313, (Albin *et al.*, 2010b)].

Finally, we noted that the 71-GxxxG-75 sequence of Vif is a common helix-helix interaction motif that facilitates interhelical hydrogen bonding, which would be consistent with an interaction between negatively charged residues of the wildtype α4 helix and either the glycine peptide backbone or the Cα-H itself as previously reported [*e.g.* (Kleiger *et al.*, 2002; Lorieau *et al.*, 2010)]. A depiction of these linear alignments is shown in **Figure 5-14A-C** for each set of APOBEC3-Vif interactions, and the helical wheel representation shows that this region of Vif, when modeled as a helix, places all of the relevant glycine residues on the same surface along with D78, which itself is aligned to APOBEC3 glycine residues. Thus, each of these residues is predicted to occur on the same helical surface and to be available for interaction with the negatively charged APOBEC3 surface (**Figure 5-14D**). We therefore hypothesized that G71D and G82D act in the context of an antiparallel helix interacting with the APOBEC3 α4 helix.

To test our hypothesis, we carried out single-cycle infectivity analyses of Vif G71D versus wildtype A3F, A3F E324K and rhA3F and of Vif G82D versus A3G and A3G D130K. As shown in **Figure 5-15A**, G71D displayed a slight loss of function in the neutralization of wildtype A3F but gained modest activity against A3F E324K and rhA3F, consistent with a role for G71 in direct interaction with A3F E324. These results were corroborated by spreading infection data demonstrating that the G71D mutation facilitates viral growth in the presence of both rhA3F and A3F QE323-324EK, albeit less robustly in the latter case, commensurate with efficiency with which G71D is selected in each condition (**Figure 5-15C-B**). At the opposite end of the putative helix-helix interaction, interpretation is somewhat complicated by the fact that the D130K change alone is only partially Vif-resistant. That said, G82D in either HIV$_{IIIB}$ or HIV$_{LAI-GFP}$

genetic backgrounds shows both enhanced degradation of A3G D130K relative to wildtype and enhanced infectivity recovery relative to wildtype (**Figure 5-15D**). Our data are therefore consistent with G71 and G82 acting as anchors on either end of a helix-helix interaction common to both A3F and A3G, indicating that the interaction of Vif with APOBEC3 proteins is, at least in this respect, structurally conserved. Analysis of additional mutants is ongoing.

Discussion

Here, we present the results of genetic selection experiments aimed at determining what residues in Vif interact with the single amino acid determinants of Vif susceptibility located in the APOBEC3 α4 helix. Despite the fact that our data suggest an antiparallel helix-helix interaction as the core of the Vif interaction with the conserved APOBEC3 α4 helix, there are several lines of evidence that suggest the overall interaction is much more complex. First, a large and growing body of literature has characterized many single amino acid changes distributed predominantly throughout the N-terminus of Vif that can render it nonfunctional against one APOBEC3 protein or another (see **Table 1-2**). While many of these, based on their sheer diversity, likely represent gross structural distortions of the protein, the broader point that many regions of Vif can affect the APOBEC3-Vif interaction is clear.

While the emergence of G71D is the clearest single result in the selection data and this mutant works very effectively at neutralizing rhA3F in spreading infections, we have found its phenotype more limited in cells expressing A3F QE323-324EK. This may, counterintuitively, be consistent with the notion that the 11 amino acids spanning 71-82 form a glycine-rich helix that interacts with the APOBEC3 α4 helix. GxxxG motifs typically mediate tight packing of the associated helices, yet this is not likely possible for the A3F QE323-324EK mutant since, even if Vif can acquire a negative charge to interact with the target lysine, the bulk of the interacting side chains would impede the broader interaction. Indeed, we have tested one molecular clone based on a selected virus that carries both G71D and G75E, which is much more efficient at spreading in A3F

187

QE323-324EK cells than G71D alone (data not shown). This may suggest the need for a longer sidechain at the other end of the GxxxG motif to stabilize the interaction.

With respect to A3G, we have found that G82D can neutralize the A3G D130K change in a single-cycle infectivity assay, suggesting a second direct interaction at the opposite end of the putative helix. The spreading infection phenotypes of this mutant are also of interest for at least two reasons. First, the G82D mutant is deficient for the neutralization of wildtype A3F [data not shown and (Dang *et al.*, 2010a)]. This may further be consistent with an interaction between G82 and D313 in A3F (the equivalent of A3G D130) since, in G82D context, these charges would repel each other. A similar loss of function phenotype does not, however, occur with G82D in the context of wildtype A3G as one might predict, indicating that, even if this putative helix represents a commonly held APOBEC3-Vif interaction motif as we propose, there must be modifications to that core interaction in different APOBEC3 proteins that result in some variation.

The G82D change was not selected using cells expressing A3G D130K, but rather DPD128-130KPK. We have not observed an effect either in single-cycle or in spreading infection to date of G82D on this double mutant, indicating that the D128K change alone must be capable of continuing to act as a dominant interrupter of the APOBEC3-Vif interaction despite apparent neutralization of A3G D130K itself. It is possible that G82D alone proffers some modest advantage to viruses in culture without itself facilitating complete adaptation to A3G DPD128-130KPK in accordance with the relatively weak selection for its occurrence observed. Indeed, we postulate that the D128K change may be the reason we observe such little selection for any sort of change in A3G, as this

188

implies any of several possibilities: 1) D128K may grossly alter the Vif interaction surface. We think this relatively unlikely since we would not likely select G82D at all if this were the case, but it is possible. 2) The effect of A3G D128K on interaction with Vif may be indirect, implicating some binding partner facilitating interaction at this residue. In particular, it is interesting to note that W127 directly adjacent to D128 is critical for A3G binding to RNA (Friew *et al.*, 2009; Huthoff *et al.*, 2009; Huthoff and Malim, 2007). Thus, there is likely RNA bound locally around D128. In the context of the wildtype residue, this may not be problematic for Vif since the negatively charged RNA backbone and D128 would repel each other and permit G84 (a fourth glycine if one extends the interaction surface past G82D and into predicted APOBEC3 loop space) to interact with D128 – or perhaps to interact with the RNA backbone itself. Placing a lysine at D128, however, would create a strong charge interaction between residue 128 and the adjacent RNA that could shield the residue from interaction with Vif. This conception is consistent with the strong loss of Vif function against both A3F and A3G associated with the G84D mutation [(Dang *et al.*, 2010a) and data not shown]. In any event, it is clear that A3G D128K strongly dominantly inhibits interaction with Vif despite extensive efforts to select a compensatory change.

In summary, we have described here the first evidence for a direct interaction between any two points in APOBEC3 proteins and Vif. Intriguingly, one may draw a line between these two points to identify a GxxxG helix-helix-interaction motif, both residues of which affect the interaction of A3F with the Vif-resistant derivative A3F QE323-324EK. We therefore propose that the core of the APOBEC3-Vif interaction is a helix-helix interaction between the α4 helix of APOBEC3 proteins that includes most of the

known single amino acid determinants of Vif susceptibility (Albin *et al.*, 2010b) and a glycine-rich helix in Vif extending from amino acids 71-82. This represents an important advance in our understand of how APOBEC3 proteins and Vif recognize each other and provides important insights for further explorations both structural and genetic of bona fide direct interactions between APOBEC3 proteins and Vif that may be targeted therapeutically.

Materials and Methods

Plasmids – Human A3F and A3G are identical to sequences NM_145298 and NM_021822, respectively, and have been previously described, as have the Vif-resistant variants A3F QE323-324EK and A3G DPD128-130KPK. Rhesus A3F and A3G were provided by Dr. Theodora Hatziioannou (Aaron Diamond AIDS Research Center, New York) and cloned into a pcDNA3.1-derived 3xHA vector as previously described (Hultquist *et al.*, 2011); the coding sequences are identical to NM_001042373.1 and AY331716.1. Note that although the equivalent of human A3G D130 is aspartate in our rhesus A3G, other variants include asparagine at that residue (*e.g.* NM001198693.1). We have utilized lysine in the rhesus-based human construct A3G DPD128-130KPK described here since lysine displays the strongest Vif-resistant phenotype at residue 130 and also allows for the clear identification of a putative compensatory charge change (Huthoff and Malim, 2007). V5-tagged derivatives of rhA3F and rhA3G were made by subcloning from the 3xHA vector as previously described (Albin *et al.*, 2010b). Specific point mutants were made by QuickChange site-directed mutagenesis (Stratagene). Full-length proviral HIV_{IIIB} and $HIV_{LAI-GFP}$ plasmids have been described previously (Albin *et al.*, 2010b); the HIV_{IIIB} sequence used is identical to that deposited as EU541617 except for an A200C nucleotide change to interrupt an aberrant upstream open reading frame in the parent virus (Haché *et al.*, 2009; Haché *et al.*, 2008). For proviral mutant plasmids, mutations were made in TOPO shuttle vectors containing the Vif-Vpr region of each provirus and then subcloned back into the full-length proviral context via the unique SwaI/SalI sites in HIV_{IIIB} or PshAI/SalI sites in $HIV_{LAI-GFP}$. All constructs were verified by sequencing and diagnostic restriction mapping.

Cell lines – The APOBEC3-expressing cell lines used in this study have been described previously (Albin *et al.*, 2010a; Albin *et al.*, 2010b; Hultquist *et al.*, 2011). In brief, SupT11 cells were electroporated with linearized pcDNA3.1-derived plasmids encoding the indicated APOBEC3 protein, and single-cell clones were isolated by limiting dilution and G418 selection for integrated constructs. Expression levels for all cells were determined by lysing $5x10^6$ cells in 250 μL of a lysis buffer consisting of 25 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM $MgCl_2$, 50 μM $ZnCl_2$, 10% glycerol and 1% Triton X-100 and supplemented with 50 μM MG132 and complete protease inhibitor (Roche), boiling portions of these lysates in a 5x sample buffer consisting of 62.5 mM Tris pH 6.8, 20% glycerol, 2% sodium dodecyl sulfate, 5% β-mercaptoethanol, and 0.05% bromophenol blue at a final concentration of 2x and fractionating proteins by SDS-PAGE. Following transfer, PVDF membranes were blocked with 4% milk in PBS-0.01% Tween, and samples were probed for A3G (#10201 from Dr. Jaisri Lingappa via the AIDS Research and Reference Reagent Program), A3F (#11474 from Dr. Michael Malim via the AIDS Research and Reference Reagent Program), hemagglutinin (HA.11, Covance) or tubulin (Covance) expression followed by appropriate secondary antibodies (anti-rabbit or anti-mouse from BioRad and Jackson, respectively). Between primary antibody probings, membranes were stripped at 50°C in a solution consisting of 62.5 mM Tris pH 6.8, 2% SDS, and 100 mM β-mercaptoethanol.

Cell culture – T cell lines were maintained in RPMI medium supplemented with 10% fetal bovine serum, penicillin/streptomycin and β-mercaptoethanol. 293T cells were

maintained in DMEM medium supplemented with 10% fetal bovine serum and penicillin/streptomycin.

Definition of adapted viruses – The selection protocol was carried out as described in the text and as depicted in **Figure 5-1**. Adapted viruses were initially selected for confirmatory passage based on any of three qualitative characteristics indicating potentially enhanced replication relative to the uninfected control well on each plate and to the characteristics of neighboring viral cultures – infectivity on CEM-GFP reporter cells, qualitatively greater syncytia formation or qualitatively enhanced cell death. An adapted phenotype in subsequent passages was ultimately defined by the following criteria: i) Adapted viruses must be blindly identified as phenotypically adapted by the aforementioned characteristics at the end of each round of selection subsequent to the first in which they were identified. ii) Adapted viruses must rise above background peak parental infectivity within 21 days during confirmatory passage on cultures composed exclusively of cells expressing the Vif-resistant APOBEC3 protein to which the given virus was initially adapted, showing at least 2% absolute CEM-GFP infectivity for viruses adapted to A3F QE323-324EK or A3G DPD128-130KPK or at least 5% for viruses adapted to rhA3F or rhA3G. iii) Adapted viruses must peak in second and third MOI-controlled confirmatory passages following the round of selection in which they were initially identified on at least one of two selective cell lines used and must remain above 1% infectivity for at least one additional time point outside of their peaks. iv) Adapted viruses must peak in a fourth confirmatory passage on at least 2/4 cell lines used within 21 days as defined in (i).

193

Spreading infections – 150,000 cells per 1 mL were plated in 24-well plates and infected with a volume of virus corresponding to the MOI in a given figure legend. Viral spread was subsequently monitored by periodic harvest of 150 μL of culture supernatant, which was then used to infect 100 μL of 25,000 CEM-GFP reporter cells in 96-well plates. Approximately 3-5 days post-infection, CEM-GFP reporter cells were fixed and subsequently analyzed by flow cytometry, with the percentage of GFP+ cells in each well taken as an indicator of infectivity. Parent cultures were periodically split and fed to prevent cell overgrowth.

Viral genotyping – To identify the *vif* sequences associated with adapted viruses, half of the CEM-GFP reporter cells infected at or near the peak of infectivity during fourth confirmatory passage were seeded into 500,000 uninfected CEM-GFP (the other half being those fixed for flow cytometry). Several days later, gDNA was harvested from these seeded cultures using the PureGene gDNA isolation method, genomic DNA was subjected to treatment with DpnI to remove any viral plasmid DNA that may have survived the many passages and manipulations to that point, and 550 ng were used as template in a touchdown PCR reaction to amplify the *vif-vpr* region as previously described (Albin *et al.*, 2010a). PCR products were visualized in ethidium bromide stained agarose gels, and bands were extracted and purified (Epoch or Fermentas gel extraction kits) for subsequent sequencing of PCR products per service instructions (Genewiz).

Single-cycle infectivity and Vif sensitivity determination – Single-cycle infectivity experiments were carried out by seeding 250,000 293 T cells in 6-well plates. The following day, these were transfected using Trans-IT transfection reagent (Mirus Bio) with 1.6 μg Vif-deficient proviral plasmid, 200 ng of an APOBEC3-V5 construct and 0, 50 or 100 ng of a Vif-HA construct. Two days post-infection, supernatants were purified through 0.45 μm filters and used to infect CEM-GFP reporter cells; three days later, these CEM-GFP reporter cells were fixed in 4% paraformaldehyde and analyzed by flow cytometry. Producer cells were also harvested at the same time as CEM-GFP infection by washing in PBS and then lysing in 250 μL lysis buffer before analysis by Western blotting as above except with anti-V5 (Invitrogen), anti-HA.11 (Covance) or anti-tubulin (Covance) antibodies.

Table 5-1: Fourth passage *vif* genotyping and phenotypic characterization of HIV$_{IIIB}$-derived viruses adapted to growth in the presence of A3F QE323-324EK.

Isolate	*vif* genotype	Growth on QE323-324EK	Growth on rhA3F	Growth on A3G DPD128-130KPK	Growth on rhA3G	Growth on CEM2n	Round First Apparent
Q1-A2	A149T, W174X	+	+	+	+	-	2
Q1-A3	E45D, M189I	++	++	+	-	+	3
Q1-B4	G71D, G75E, V142I	++	++	+	-	-	3
Q1-B5	I31V, G191R*	++	++	+	+	+	2
Q1-C1	nd	+	+	+	-	-	2
Q1-C2	K26N, E45D,	++	+	+	-	-	1
Q1-C3	G71D,	+	nd	nd	nd	nd	3
Q1-C4	T180N	+	nd	nd	nd	nd	3
Q1-C5	I31V	++	++	+	-	-	2
Q1-D1	H43Y, K50R, E117K	++	+	-	-	-	1
Q3-A1	E45D, s116t,	++	+	+	-	-	3
Q3-B6	G71D, M189I, G191E*	++	+++	+	+	+	3

Criteria

- = No growth; for rhA3G, peaks one time point or later than vector

+ = Weak growth such as failure to peak on one of the cell lines expressing A3F QE323-324EK or delayed peaks in the last two time points (~3 weeks); for rhA3G, peaks at or before vector

++ = Growth with peaks at the middle two time points (~2 weeks); no criteria for rhA3G

+++ = Growth with peaks within the first two time points (~1 week); no criteria for rhA3G

nd = Not determined, typically due to shortage of isolate viruses with which to initiate infection

*Indicates mutation that introduces a stop codon in the overlapping *vpr* open reading frame

Lower case *vif* genotypes indicate a minor population

Table 5-2: Fourth passage *vif* genotyping and phenotypic characterization of HIV$_{\text{LAI-GFP}}$-derived viruses adapted to growth in the presence of A3F QE323-324EK.

Isolate	*vif* genotype	Growth on QE323-324EK	Growth on rhA3F	Growth on A3G DPD128-130KPK	Growth on rhA3G	Growth on CEM2n	Round First Apparent
Q4-A1	G71D, s116t, K160R	++	++	+	-	na	2
Q4-A2	D117N	++	nd	nd	nd	na	2
Q4-A3	nd	+	-	-	-	na	3
Q4-A4	RL63-64KM, G191E*	++	++	-	-	na	2
Q4-B3	WT	+	nd	nd	nd	na	1
Q4-B4	D117N, L153I	++	++	+	+	na	1
Q4-C1	k92r, G185R	++	++	-	-	na	2
Q4-C2	G185R	++	++	+	-	na	3
Q4-C3	k92r	++	++	-	+	na	1
Q4-C4	WT	++	++	++	+	na	1
Q4-C5	nd	++	++	+	+	na	1
Q4-C6	WT	++	++	++	+	na	3
Q4-D1	G71D, G185R	++	++	+	+	na	1
Q4-D2	R19K	++	++	+	-	na	2
Q4-D4	k92r, M189K	++	+++	++	+	na	3
Q11-A1	WT	++	++	-	-	na	2
Q11-A2	R19K, K26E	++	++	-	+	na	1
Q11-A3	WT	+	++	-	+	na	2
Q11-A5	K92R	++	++	+	-	na	3
Q11-A6	g71d, G185R	++	++	-	-	na	1
Q11-B1	L153I	++	++	-	-	na	1
Q11-B2	WT	++	++	+	-	na	3

Isolate	*vif* genotype	Growth on QE323-324EK	Growth on rhA3F	Growth on A3G DPD128-130KPK	Growth on rhA3G	Growth on CEM2n	Round First Apparent
Q11-B3	g71d	++	++	+	-	na	3
Q11-B4	k141r	++	++	++	+	na	2
Q11-B5	k92r, n175k, M189I	++	++	-	nd	na	2
Q11-B6	WT	+	++	-	+	na	2
Q11-C1	D117N	++	++	+	+	na	2
Q11-C2	v98i, G191R*	++	++	+	-	na	1
Q11-C3	G191R*	++	++	+	+	na	3
Q11-C4	WT	++	++	-	-	na	2
Q11-C5	WT	++	++	+	-	na	1
Q11-C6	WT	++	++	+	-	na	1
Q11-D1	D117N, G185R	++	++	-	-	na	2
Q11-D2	g71d, g191e*	++	++	+	nd	na	1
Q11-D3	WT	++	++	-	-	na	1
Q11-D4	nd	++	nd	nd	nd	na	1

Criteria

- = No growth; for rhA3G, peaks one time point or later than vector

+ = Weak growth such as failure to peak on one of the cell lines expressing A3F QE323-324EK or delayed peaks in the last two time points (~3 weeks); for rhA3G, peaks at or before vector

++ = Growth with peaks at the middle two time points (~2 weeks); no criteria for rhA3G

+++ = Growth with peaks within the first two time points (~1 week); no criteria for rhA3G

nd = Not determined, typically due to shortage of isolate viruses with which to initiate infection

na = Not applicable; HIV$_{LAI-GFP}$ derivatives do not grow on CEM2n cells due to interruption of the *nef* open reading frame by GFP [*e.g.* (Chowers *et al.*, 1994)].
*Indicates mutation that introduces a stop codon in the overlapping *vpr* open reading frame
Lower case *vif* genotypes indicate a minor population

Table 5-3: Fourth passage *vif* genotyping and phenotypic characterization of HIV_{IIIB}-derived viruses adapted to growth in the presence of A3G DPD128-130KPK.

Isolate	*vif* genotype	Growth on A3G DPD128-130KPK	Growth on rhA3G	Growth on QE323-324EK	Growth on rhA3F	Growth on CEM2n	Round First Apparent
D4-B2	G71D	++	+	++	++	-	3
D4-B6	G71D, G82D	++	+	+	++	-	3
D4-C2	D78E, L163S	++	-	+	+	-	2
D4-D2	S32A, frameshift 178	++	-	+	+	-	3
D4-D3	L64S, W174X	++	-	+	+	-	2
D9-B2	K92R, P131L	++	+	+	+	+	3
D9-B3	W79X, G84X	++	+	+	+	-	2
D9-B4	frameshift 189	++	-	+	+	+	3
D9-B5	F112I	++	+	+	+	-	3
D9-D2	C114R	++	-	+	+	-	3

Criteria

- = No growth; for rhA3G, peaks one time point or later than vector

+ = Weak growth such as failure to peak on one of the cell lines expressing A3G DPD128-130KPK or delayed peaks in the last two time points (~3 weeks); for rhA3G, peaks at or before vector

++ = Growth with peaks at the middle two time points (~2 weeks); no criteria for rhA3G

+++ = Growth with peaks within the first two time points (~1 week); no criteria for rhA3G

nd = Not determined, typically due to shortage of isolate viruses with which to initiate infection

*Indicates mutation that introduces a stop codon in the overlapping *vpr* open reading frame

Minor populations not assessed in these sequences.

Table 5-4: Fourth passage *vif* genotyping and phenotypic characterization of HIV_{LAI-GFP}-derived viruses adapted to growth in the presence of A3G DPD128-130KPK.

Isolate	*vif* genotype	Growth on A3G DPD128-130KPK	Growth on rhA3G	Growth on QE323-324EK	Growth on rhA3F	Growth on CEM2n	Round First Apparent
D14-A1	P100S	++	-	+	++	na	3
D14-A3	G82D, G191R*	++	+	++	++	na	3
D14-B5	WT	+	nd	+	++	na	3
D14-C4	H127N	++	+	++	++	na	3
D14-C5	nd	+	+	+	++	na	3
D14-C6	WT	++	+	+++	+++	na	3
D15-A2	R15G, R50K, G82D, D117N	++	nd	+	+	na	3
D15-B1	WT	+	nd	++	nd	na	3
D15-B2	R15W	++	+	++	++	na	3
D15-B3	R50K, E101K	++	nd	++	++	na	3
D15-B4	WT	+	nd	+	++	na	3
D15-B6	P100S	++	+	++	++	na	3
D15-C1	WT	+	nd	nd	nd	na	3
D15-C3	WT	++	nd	nd	nd	na	3
D15-C5	nd	+	nd	nd	nd	na	3
D15-D1	R121K	++	nd	+	nd	na	3
D15-D2	H80Y	++	+	++	++	na	3
D15-D3	G60R, G191E*	+	nd	+	nd	na	3

Criteria
- = No growth; for rhA3G, peaks one time point or later than vector
+ = Weak growth such as failure to peak on one of the cell lines expressing A3G
DPD128-130KPK or delayed peaks in the last two time points (~3 weeks); for rhA3G,
peaks at or before vector
++ = Growth with peaks at the middle two time points (~2 weeks); no criteria for rhA3G
+++ = Growth with peaks within the first two time points (~1 week); no criteria for
rhA3G
nd = Not determined, typically due to shortage of isolate viruses with which to initiate
infection
na = Not applicable; HIV$_{LAI-GFP}$ derivatives do not grow on CEM2n cells due to
interruption of the *nef* open reading frame by GFP [*e.g.* (Chowers *et al.*, 1994)].
*Indicates mutation that introduces a stop codon in the overlapping *vpr* open reading
frame
Minor populations not assessed in these sequences.

Table 5-5: Fourth passage *vif* genotyping and phenotypic characterization of HIV_{IIIB}-derived viruses adapted to growth in the presence of rhA3F.

Isolate	*vif* Genotype	Growth on rhA3F	Growth on A3F QE323-324EK	Round First Apparent
50-4(1)-A1	G71D, T74I	++	+	3
50-4(1)-A6	G71D	+++	+++	2
50-4(1)-B1	nd	nd	nd	3
50-4(1)-B2	G71D	+++	++	2
50-4(1)-B3	G71D	+++	+++	3
50-4(1)-B5	G71D, E117K	+++	++	1
50-4(1)-C1	nd	++	++	2
50-4(1)-C2	R23S	++	-	3
50-4(1)-C5	G71D	+++	++	2
50-4(1)-D1	G71D	+++	+++	3
50-4(1)-D2	G71D, Q83R	+++	+++	3
50-4(1)-D3	G71D	+++	+++	2
50-4(1)-D4	G71D	+++	+++	2
50-4(2)-A1	G71D, I128R	++	++	2
50-4(1)-A4	G71D, Y111H, G191R*	+++	++	2
50-4(2)-A5	L59V, G71D, R127C	++	++	3
50-4(2)-A6	G71D, E117K	+++	++	3
50-4(2)-B3	G71D, L81W	+++	++	2
50-4(2)-B5	G71D	+++	+++	1
50-4(2)-B6	G71D	+++	++	2
50-4(2)-C2	R23S, I31V, E117K	+++	+	3
50-4(2)-C3	nd	+++	++	2
50-4(2)-C4	G71D	+++	++	3
50-4(2)-C5	S116L, I12L	++	+	3
50-4(2)-C6	nd	+++	++	2
50-4(2)-D1	nd	+++	++	3
50-4(2)-D3	nd	+	-	3

Criteria

- = No growth

+ = Weak growth such as failure to peak on one of the selective cell lines used or late peaks within the last two time points.

++ = Growth with peaks by Day 13 post-infection

+++ = Growth with peaks by Day 8 post-infection

nd = Not determined, typically due to shortage of isolate viruses with which to initiate infection.

Minor populations not assessed in these sequences.

Table 5-6: Fourth passage *vif* genotyping and phenotypic characterization of HIV_{LAI-GFP}-derived viruses adapted to growth in the presence of rhA3F.

Isolate	*vif* Genotype	Growth on rhA3F	Growth on A3F QE323-324EK	Round First Apparent
50-5(1)-A1	nd	+++	++	2
50-5(1)-A3	nd	++	++	3
50-5(1)-A4	nd	++	++	2
50-5(1)-A5	D117N	++	++	1
50-5(1)-A6	WT	+	-	2
50-5(1)-B1	WT	++	+	1
50-5(1)-B4	WT	++	++	1
50-5(1)-B6	nd	++	++	3
50-5(1)-C1	G191E*	++	++	1
50-5(1)-C4	nd	++	++	2
50-5(1)-C5	nd	++	+	3
50-5(1)-D1	nd	++	++	2
50-5(1)-D2	I159M	++	++	2
50-5(1)-D3	WT	++	++	1
50-5(1)-D4	K92R	++	++	2
50-5(2)-A1	WT	++	++	2
50-5(2)-A2	R93K	++	++	2
50-5(2)-A4	WT	++	++	3
50-5(2)-A5	WT	++	+	3
50-5(2)-B1	nd	++	++	1
50-5(2)-B2	nd	++	++	2
50-5(2)-B3	WT	++	++	3
50-5(2)-B4	G191E*	++	++	2
50-5(2)-B5	G71D, D78E	++	++	2
50-5(2)-B6	WT	++	++	2
50-5(2)-C1	WT	++	++	2
50-5(2)-C2	nd	++	++	3
50-5(2)-C3	nd	++	+	3
50-5(2)-C4	nd	++	++	2
50-5(2)-C5	nd	nd	nd	1
50-5(2)-D1	nd	nd	nd	2

Criteria

- = No growth

+ = Weak growth such as failure to peak on one of the selective cell lines used or late peaks within the last two time points.

++ = Growth with peaks by Day 13 post-infection

+++ = Growth with peaks by Day 8 post-infection

nd = Not determined, typically due to shortage of viruses with which to initiate infection. Minor populations not assessed in these sequences.

Table 5-7: Fourth passage *vif* genotyping and phenotypic characterization of HIV_{IIIB}-derived viruses adapted to growth in the presence of rhA3G.

Isolate	*vif* Genotype	Growth on rhA3G	Growth on A3G DPD/KPK	Round First Apparent
51-3(1)-A2	nd	++	-	2
51-3(1)-A3	nd	+++	++	1
51-3(1)-A6	nd	nd	nd	2
51-3(1)-A5	nd	nd	nd	3
51-3(1)-B1	M8I, G71D	+++	+	3
51-3(1)-B2	R33K, G71D	+++	+	2
51-3(1)-B3	nd	nd	nd	1
51-3(1)-B4	nd	nd	nd	3
51-3(1)-B5	I18S, T188K*	++	+	3
51-3(1)-B6	V98I	++	-	3
51-3(1)-C6	I31S	+++	+	3
51-3(1)-D4	T67I, S95N	+++	-	3
51-3(2)-A1	K50N	++	-	3
51-3(2)-A4	R17N	nd	nd	3
51-3(2)-A6	M29I	++	-	3
51-3(2)-B3	nd	++	+	1
51-3(2)-B4	nd	++	+	1
51-3(2)-B5	nd	+++	-	2
51-3(2)-B6	V85A	nd	nd	3
51-3(2)-C3	nd	nd	nd	3
51-3(2)-C4	I51T, L59Q	+++	-	1
51-3(2)-C5	WT	++	-	3
51-3(2)-C6	WT	++	+	2
51-3(2)-D2	WT	++	-	2
51-3(2)-D3	nd	++	-	2

Criteria

- = No growth

+ = Weak growth such as failure to peak on one of the selective cell lines used or late peaks within the last two time points.

++ = Growth with peaks by Day 13 post-infection

+++ = Growth with peaks by Day 8 post-infection

nd = Not determined, typically due to shortage of isolate viruses with which to initiate infection.

Minor populations not assessed in these sequences.

Table 5-8: Fourth passage *vif* genotyping and phenotypic characterization of HIV$_{LAI-GFP}$-derived viruses adapted to growth in the presence of rhA3G.

Isolate	*vif* Genotype	Growth on rhA3G	Growth on DPD/KPK	Round First Apparent
51-5(1)-A2	WT	+++	+	3
51-5(1)-A3	Q136R	+	-	2
51-5(1)-A4	nd	+++	-	1
51-5(1)-A5	nd	+	-	1
51-5(1)-A6	nd	+++	+	2
51-5(1)-B1	D117V	++	+	3
51-5(1)-C1	R4K	++	-	2
51-5(1)-C2	R15G, R50K, G82D, D117N	++	++	2
51-5(1)-C3	WT	+++	+	2
51-5(1)-C4	WT	++	+	2
51-5(1)-C6	WT	+++	++	1
51-5(1)-D1	WT	++	+	1
51-5(1)-D2	WT	+++	++	2
51-5(1)-D4	WT	++	+	3
51-5(2)-A4	L59I	++	+	3
51-5(2)-A5	WT	++	+	2
51-5(2)-B2	G185R	++	+	2
51-5(2)-B4	nd	++	-	2
51-5(2)-B5	WT	+	-	2
51-5(2)-B6	nd	++	-	2
51-5(2)-C1	WT	+++	-	3
51-5(2)-C5	nd	+	-	3
51-5(2)-C6	WT	+	-	1
51-5(2)-D1	nd	++	-	1
51-5(2)-D3	nd	nd	nd	3

Criteria

- = No growth

+ = Weak growth such as failure to peak on one of the selective cell lines used or late peaks within the last two time points.

++ = Growth with peaks by Day 13 post-infection

+++ = Growth with peaks by Day 8 post-infection

nd = Not determined, typically due to shortage of isolate viruses with which to initiate infection.

Minor populations not assessed in these sequences.
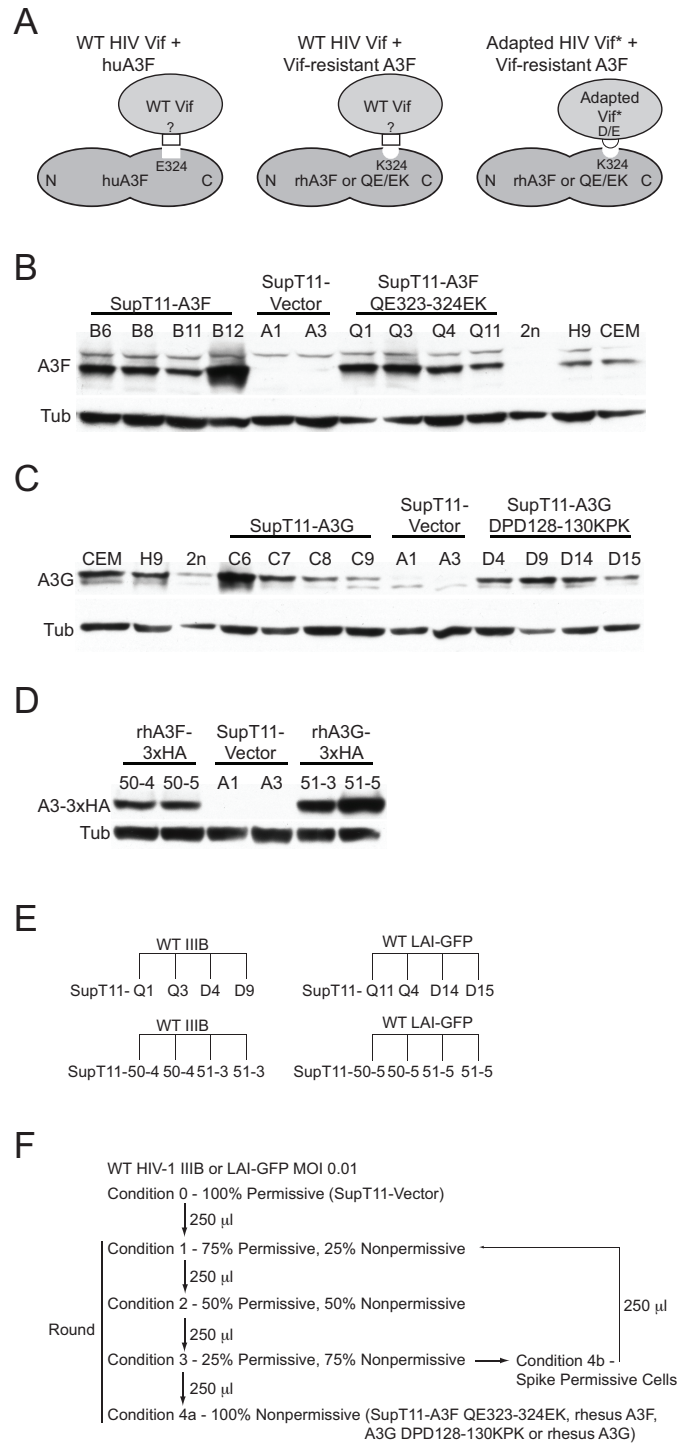
Figure 5-1: Experimental approach to the genetic identification of residues in Vif that directly interact with the APOBEC3 α4 helix determinants of Vif susceptibility.

Figure 5-1: Experimental approach to the genetic identification of residues in Vif that directly interact with the APOBEC3 α4 helix determinants of Vif susceptibility.

(A) A conceptual depiction of the experimental strategy using A3F as an example. Wildtype HIV Vif interacts with wildtype human A3F in the left panel, permitting viral spread. In the middle panel, a negative-to-positive charge change E324K renders A3F Vif-resistant, preventing viral spread. At the right, a compensatory negative charge in Vif restores functional interaction with A3F E324K, reversing the Vif-resistant phenotype and once again permitting viral spread. **(B)** SupT11-derived T cell lines used in this study that have been stably transfected with human A3F, A3F QE323-324EK, or a vector control as previously described (Albin *et al.*, 2010b; Refsland *et al.*, 2010). **(C)** SupT11-derived T cell lines used in this study that have been stably transfected with human A3G, A3G DPD128-130KPK or a vector control, as previously described (Albin *et al.*, 2010b; Refsland *et al.*, 2010). **(D)** SupT11-derived T cell lines used in this study that have been stably transfected with rhesus macaque A3F-3xHA, rhesus macaque A3G 3x-HA or a vector control, as previously described (Hultquist *et al.*, 2011). **(E)** A schematic showing the selection conditions used in this study in which each prong represents a 24-well plate with 22 selective cultures, an infected permissive cell culture and an uninfected culture. **(F)** A schematic depicting the co-culture selection strategy used in this chapter. Approximately 8-11 days separate each passage from the next. A "Round" of selection refers collectively to a set of step-wise passages from permissive to increasingly nonpermissive cultures as shown. Viruses were evaluated at Condition 4(a) for potential MOI-controlled, confirmatory passages, while portions of each culture were also cycled

back to the beginning of the process for another Round of selection as shown in Condition 4(b).
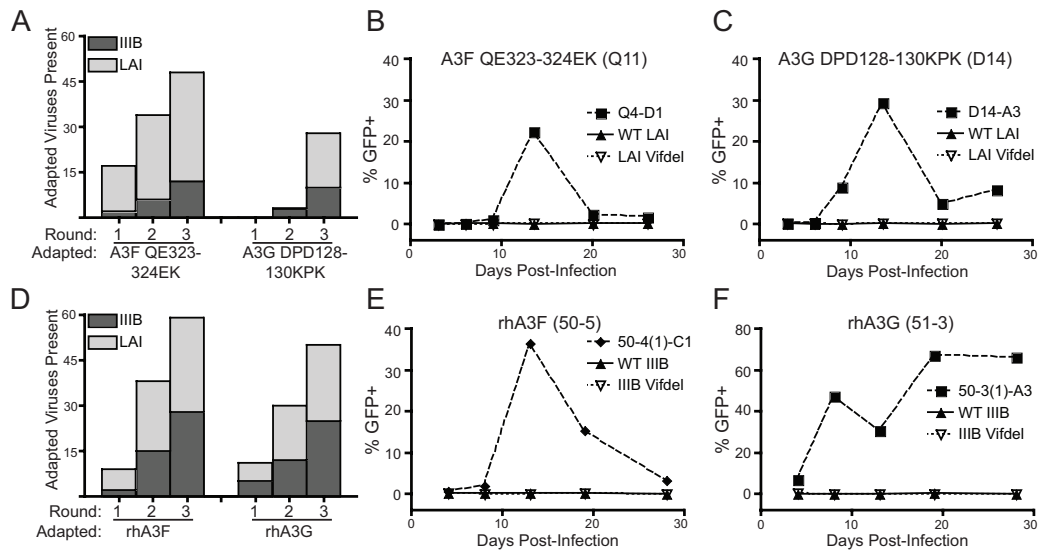
Figure 5-2: Successful selection for adapted viruses. (A) Histograms showing the total number of adapted isolates present at the end of each round of selection as determined in **Materials and Methods** for A3F QE323-324EK and A3G DPD128-130KPK. **(B-C)** Representative adapted viruses from the fourth passage of the A3F QE323-324EK and A3G DPD128-130KPK adapted viruses are shown relative to the parental molecular clone derived viruses and Vif-deficient derivatives of the same molecular clones. **(D-F)** As in **A-B** but with rhA3F and rhA3G.
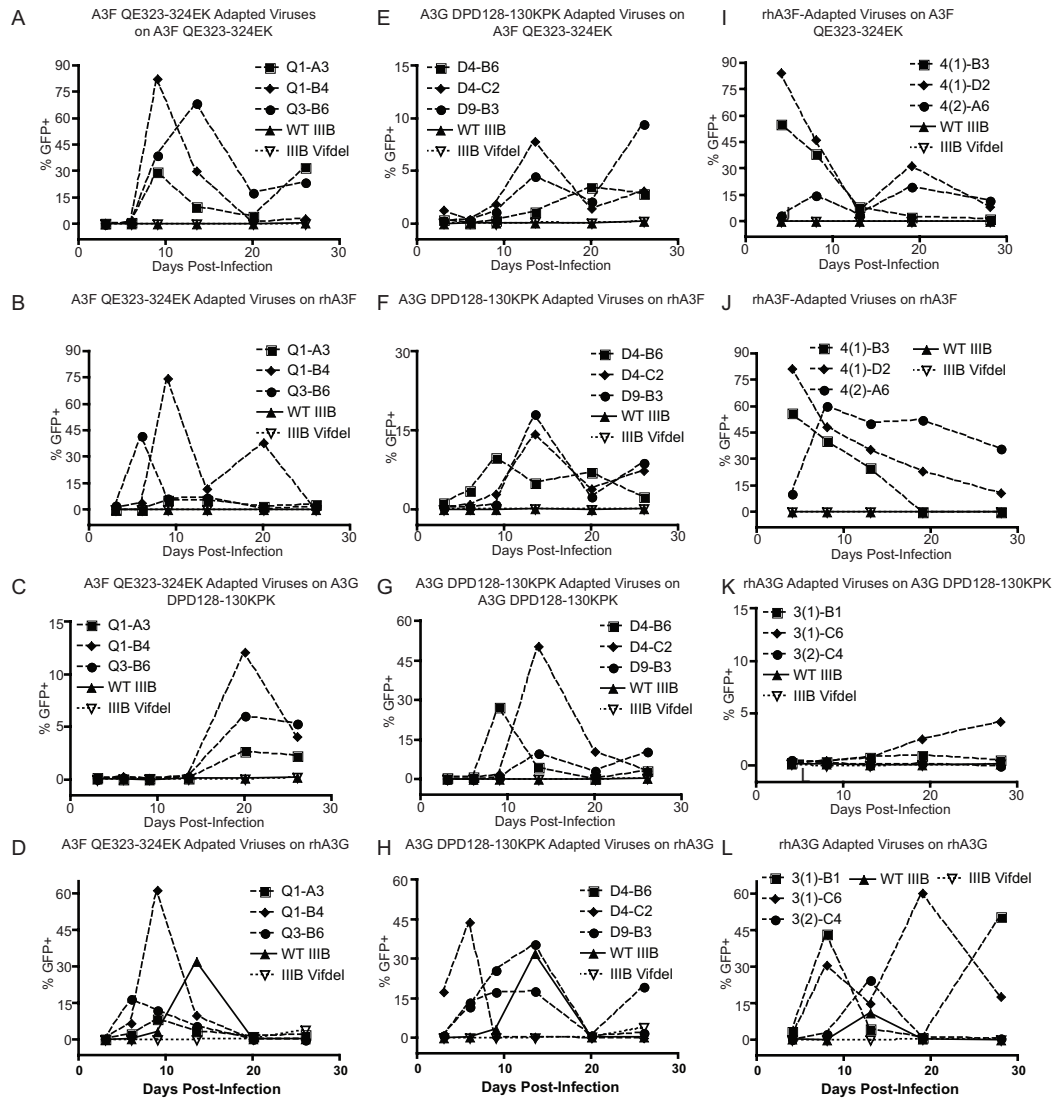
Figure 5-3: Representative native and cross-resistance phenotypes of adapted viruses.

211

Figure 5-3: Representative native and cross-resistance phenotypes of adapted viruses. (A-D) Representative isolates originally adapted to A3F QE323-324EK and their growth on both a native adaptive cell line **(A)** and on cell lines expressing **(B)** rhA3F, **(C)** A3G DPD128-130KPK and **(D)** rhA3G. **(E-H)** Representative isolates originally adapted to A3G DPD128-130KPK and their growth on both a native adaptive cell lines **(G)** and on cell lines expressing **(H)** rhA3G, **(E)** A3F QE323-324EK and **(F)** rhA3F. **(I-J)** Representative isolates originally adapted to rhA3F and their growth on both a native adaptive cell line **(J)** and on cells expressing **(I)** A3F QE323-324EK. **(K-L)** Representative isolates originally adapted to rhA3G and their growth on both a native adaptive cell line **(L)** and on cells expressing **(K)** A3G DPD128-130KPK. Note that the fourth passages of viruses adapted to A3F QE323-324EK and A3G DPD128-130KPK were carried out as one experiment from MOI 0.02, while the fourth passage of viruses adapted to rhA3F or rhA3G were carried out as another, independent experiment from MOI 0.02. Thus, the control wildtype and Vif-deficient viruses replotted for the same cell line in the first two columns (*e.g.* panels **A** and **E**) are the same.

```
IIIB_Vif    MENRWQVMIV WQVDRMRINT WKRLVKHHMY ISRKAKDWFY RHHYESTNPK ISSEVHIPLG DAKLVITTYW
Q1-A2       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-A3       ---------- ---------- ---------- ---------- ----D----- ---------- ----------
Q1-B4       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-B5       ---------- ---------- ---------- V--------- ---------- ---------- ----------
Q1-C2       ---------- ---------- -----N---- ---------- ----D----- ---------- ----------
Q1-C3       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-C4       ---------- ---------- ---------- ---------- ----K----- ---------- ----------
Q1-C5       ---------- ---------- ---------- V--------- ---------- ---------- ----------
Q1-D1       ---------- ---------- ---------- ---------- --Y------R ---------- ----------
Q3-A1       ---------- ---------- ---------- ---------- ----D----- ---------- ----------
Q3-B6       ---------- ---------- ---------- ---------- ---------- ---------- ----------   70


IIIB_Vif    GLHTGERDWH LGQGVSIEWR KKRYSTQVDP DLADQLIHLH YFDCFSESAI RNTILGRIVS PRCEYQAGHN
Q1-A2       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-A3       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-B4       D---E----- ---------- ---------- ---------- ---------- ---------- ----------
Q1-B5       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-C2       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-C3       D--------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-C4       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-C5       ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q1-D1       ---------- ---------- ---------- ---------- ------K--- ---------- ----------
Q3-A1       ---------- ---------- ---------- ---------- -----t---- ---------- ----------
Q3-B6       D--------- ---------- ---------- ---------- ---------- ---------- ----------   140


IIIB_Vif    KVGSLQYLAL AALIKPKQIK PPLPSVRKLT EDRWNKPQKT KGHRGSHTMN GH*
Q1-A2       --------T- ---------- ---------- ---*------ ---------- ---
Q1-A3       ---------- ---------- ---------- ---------- --------I- ---
Q1-B4       -I-------- ---------- ---------- ---------- ---------- ---
Q1-B5       ---------- ---------- ---------- ---------- ---------- R--
Q1-C2       ---------- ---------- ---------- ---------- ---------- ---
Q1-C3       ---------- ---------- ---------- --------N ---------- ---
Q1-C4       ---------- ---------- ---------- ---------- ---------- ---
Q1-C5       ---------- ---------- ---------- ---------- ---------- ---
Q1-D1       ---------- ---------- ---------- ---------- ---------- ---
Q3-A1       ---------- ---------- ---------- ---------- ---------- ---
Q3-B6       ---------- ---------- ---------- ---------- --------I- E--      192
```

Figure 5-4: Alignment of HIV$_{IIIB}$ *vif* alleles selected by A3F QE323-324EK. All

alignments prepared using SeqPublish: http://www.hiv.lanl.gov/content/sequence/

SeqPublish/seqpublish.html

```
LAI_Vif    MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW
Q4-A1      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-A2      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-A4      ---------- ---------- ---------- ---------- ---------- ---------- --KM------
Q4-A6      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-B3      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-B4      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C1      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C2      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C3      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C4      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C6      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-D1      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-D2      ---------- -------K- ---------- ---------- ---------- ---------- ----------
Q4-D4      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A2     ---------- -------K- ----E---- ---------- ---------- ---------- ----------
Q11-A3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-D1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-D2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-D3     ---------- ---------- ---------- ---------- ---------- ---------- ----------      70


LAI_Vif    GLHTGERDWH LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN
Q4-A1      D--------- ---------- ---------- ---------- -----t---- ---------- ----------
Q4-A2      ---------- ---------- ---------- ---------- ------N--- ---------- ----------
Q4-A4      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-A6      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-B3      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-B4      ---------- ---------- ---------- ---------- ------N--- ---------- ----------
Q4-C1      ---------- ---------- -r-------- ---------- ---------- ---------- ----------
Q4-C2      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C3      ---------- ---------- -r-------- ---------- ---------- ---------- ----------
Q4-C4      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-C6      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-D1      D--------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-D2      ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q4-D4      ---------- ---------- -r-------- ---------- ---------- ---------- ----------
Q11-A1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A3     ---------- ---------- -R-------- ---------- ---------- ---------- ----------
Q11-A5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-A6     d--------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B3     d--------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-B5     ---------- ---------- -r-------- ---------- ---------- ---------- ----------
Q11-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C1     ---------- ---------- ---------- ---------- ------N--- ---------- ----------
Q11-C2     ---------- ---------- -------i-- ---------- ---------- ---------- ----------
Q11-C3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-D1     ---------- ---------- ---------- ---------- ------N--- ---------- ----------
Q11-D2     d--------- ---------- ---------- ---------- ---------- ---------- ----------
Q11-D3     ---------- ---------- ---------- ---------- ---------- ---------- ----------      140
```

Figure 5-5: Alignment of HIV$_{LAI-GFP}$ *vif* alleles selected by A3F QE323-324EK.

```
LAI_Vif     KVGSLQYLAL AALITPKKIK PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH*
Q4-A1       ---------- ---------R ---------- ---------- ---------- ---
Q4-A2       ---------- ---------- ---------- ---------- ---------- ---
Q4-A4       ---------- ---------- ---------- ---------- ---------- E--
Q4-A6       ---------- ---------- ---------- ---------- ---------- ---
Q4-B3       ---------- ---------- ---------- ---------- ---------- ---
Q4-B4       ---------- --I------- ---------- ---------- ---------- ---
Q4-C1       ---------- ---------- ---------- ---------- ----R----- ---
Q4-C2       ---------- ---------- ---------- ---------- ----R----- ---
Q4-C3       ---------- ---------- ---------- ---------- ---------- ---
Q4-C4       ---------- ---------- ---------- ---------- ---------- ---
Q4-C6       ---------- ---------- ---------- ---------- ---------- ---
Q4-D1       ---------- ---------- ---------- ---------- ----R----- ---
Q4-D2       ---------- ---------- ---------- ---------- ---------- ---
Q4-D4       ---------- ---------- ---------- ---------- --------K- ---
Q11-A1      ---------- ---------- ---------- ---------- ---------- ---
Q11-A2      ---------- ---------- ---------- ---------- ---------- ---
Q11-A3      ---------- ---------- ---------- ---------- ---------- ---
Q11-A5      ---------- ---------- ---------- ---------- ---------- ---
Q11-A6      ---------- ---------- ---------- ---------- ----R----- ---
Q11-B1      ---------- --I------- ---------- ---------- ---------- ---
Q11-B2      ---------- ---------- ---------- ---------- ---------- ---
Q11-B3      ---------- ---------- ---------- ---------- ---------- ---
Q11-B4      r--------- ---------- ---------- ---------- ---------- ---
Q11-B5      ---------- ---------- ---------- ----k----- --------i- ---
Q11-B6      ---------- ---------- ---------- ---------- ---------- ---
Q11-C1      ---------- ---------- ---------- ---------- ---------- ---
Q11-C2      ---------- ---------- ---------- ---------- ---------- R--
Q11-C3      ---------- ---------- ---------- ---------- ---------- R--
Q11-C4      ---------- ---------- ---------- ---------- ---------- ---
Q11-C5      ---------- ---------- ---------- ---------- ---------- ---
Q11-C6      ---------- ---------- ---------- ---------- ---------- ---
Q11-D1      ---------- ---------- ---------- ---------- ----R----- ---
Q11-D2      ---------- ---------- ---------- ---------- ---------- e--
Q11-D3      ---------- ---------- ---------- ---------- ---------- --- 192
```

Figure 5-5 (continued): Alignment of HIV_{LAI-GFP} *vif* alleles selected by A3F QE323-324EK.

```
IIIB_Vif     MENRWQVMIV WQVDRMRINT WKRLVKHHMY ISRKAKDWFY RHHYESTNPK ISSEVHIPLG DAKLVITTYW
D4-B2        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D4-B6        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D4-C2        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D4-D2        ---------- ---------- ---------- -A-------- ---------- ---------- ----------
D4-D3        ---------- ---------- ---------- ---------- ---------- ---------- ---S------
D9-B2        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D9-B3        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D9-B4        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D9-B5        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D9-D2        ---------- ---------- ---------- ---------- ---------- ---------- ----------  70

IIIB_Vif     GLHTGERDWH LGQGVSIEWR KKRYSTQVDP DLADQLIHLH YFDCFSESAI RNTILGRIVS PRCEYQAGHN
D4-B2        D--------- ---------- ---------- ---------- ---------- ---------- ----------
D4-B6        D--------- -D-------- ---------- ---------- ---------- ---------- ----------
D4-C2        -------E-- ---------- ---------- ---------- ---------- ---------- ----------
D4-D2        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D4-D3        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D9-B2        ---------- ---------- -R-------- ---------- ---------- ---------- L---------
D9-B3        --------*- ---*------ ---------- ---------- ---------- ---------- ----------
D9-B4        ---------- ---------- ---------- ---------- ---------- ---------- ----------
D9-B5        ---------- ---------- ---------- ---------- -I-------- ---------- ----------
D9-D2        ---------- ---------- ---------- ---------- ---R------ ---------- ----------  140

IIIB_Vif     KVGSLQYLAL AALIKPKQIK PPLPSVRKLT EDRWNKPQKT KGHRGSHTMN GH*
D4-B2        ---------- ---------- ---------- ---------- ---------- ---
D4-B6        ---------- ---------- ---------- ---------- ---------- ---
D4-C2        ---------- ---------- --S------- ---------- ---------- ---
D4-D2        ---------- ---------- ---------- -------RRP RATE-AIQ*M DT
D4-D3        ---------- ---------- ---------- ---*------ ---------- ---
D9-B2        ---------- ---------- ---------- ---------- ---------- ---
D9-B3        ---------- ---------- ---------- ---------- ---------- ---
D9-B4        ---------- ---------- ---------- ---------- --------IE SEA
D9-B5        ---------- ---------- ---------- ---------- ---------- ---
D9-D2        ---------- ---------- ---------- ---------- ---------- ---     192
```

Figure 5-6: Alignment of HIV_{IIIB} *vif* alleles selected by A3G DPD128-130KPK.

```
LAI_Vif    MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW
D14-A1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D14-A3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D14-B5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D14-C4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D14-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-A2     ---------- ----G----- ---------- ---------- ---------K ---------- ----------
D15-B1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-B2     ---------- ----W----- ---------- ---------- ---------- ---------- ----------
D15-B3     ---------- ---------- ---------- ---------- ---------K ---------- ----------
D15-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-C3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-D1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-D2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-D3     ---------- ---------- ---------- ---------- ---------- --------R ----------    70


LAI_Vif    GLHTGERDWH LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN
D14-A1     ---------- ---------- --------S ---------- ---------- ---------- ----------
D14-A3     ---------- -D-------- ---------- ---------- ---------- ---------- ----------
D14-B5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D14-C4     ---------- ---------- ---------- ---------- ---------- ------N--- ----------
D14-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-A2     ---------- -D-------- ---------- ---------- ------N--- ---------- ----------
D15-B1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-B2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-B3     ---------- ---------- ---------- K--------- ---------- ---------- ----------
D15-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-B6     ---------- ---------- --------S ---------- ---------- ---------- ----------
D15-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-C3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
D15-D1     ---------- ---------- ---------- ---------- ---------- K--------- ----------
D15-D2     --------Y ---------- ---------- ---------- ---------- ---------- ----------
D15-D3     ---------- ---------- ---------- ---------- ---------- ---------- ----------    140


LAI_Vif    KVGSLQYLAL AALITPKKIK PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH*
D14-A1     ---------- ---------- ---------- ---------- ---------- ---
D14-A3     ---------- ---------- ---------- ---------- ---------- R--
D14-B5     ---------- ---------- ---------- ---------- ---------- ---
D14-C4     ---------- ---------- ---------- ---------- ---------- ---
D14-C6     ---------- ---------- ---------- ---------- ---------- ---
D15-A2     ---------- ---------- ---------- ---------- ---------- ---
D15-B1     ---------- ---------- ---------- ---------- ---------- ---
D15-B2     ---------- ---------- ---------- ---------- ---------- ---
D15-B3     ---------- ---------- ---------- ---------- ---------- ---
D15-B4     ---------- ---------- ---------- ---------- ---------- ---
D15-B6     ---------- ---------- ---------- ---------- ---------- ---
D15-C1     ---------- ---------- ---------- ---------- ---------- ---
D15-C3     ---------- ---------- ---------- ---------- ---------- ---
D15-D1     ---------- ---------- ---------- ---------- ---------- ---
D15-D2     ---------- ---------- ---------- ---------- ---------- ---
D15-D3     ---------- ---------- ---------- ---------- ---------- E-- 192
```

Figure 5-7: Alignment of HIV_{LAI-GFP} *vif* alleles selected by A3G DPD128-130KPK.

```
IIIB_Vif     MENRWQVMIV WQVDRMRINT WKRLVKHHMY ISRKAKDWFY RHHYESTNPK ISSEVHIPLG DAKLVITTYW
50-4(1)-A1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-A6   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-B2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-B3   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-B5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-C2   ---------- ---------- --S------- ---------- ---------- ---------- ----------
50-4(1)-C5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D3   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-A1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-A4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-A5   ---------- ---------- ---------- ---------- ---------- -------V- ----------
50-4(2)-A6   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-B3   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-B5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-B6   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-C2   ---------- ---------- --S------- V--------- ---------- ---------- ----------
50-4(2)-C4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-C5   ---------- ---------- ---------- ---------- ---------- ---------- ----------   70


IIIB_Vif     GLHTGERDWH LGQGVSIEWR KKRYSTQVDP DLADQLIHLH YFDCFSESAI RNTILGRIVS PRCEYQAGHN
50-4(1)-A1   D--I------ ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-A6   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-B2   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-B3   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-B5   D--------- ---------- ---------- ---------- ------K--- ---------- ----------
50-4(1)-C2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-C5   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D1   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D2   D--------- --R------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D3   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(1)-D4   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-A1   D--------- ---------- ---------- ---------- ---------- -------R-- ----------
50-4(2)-A4   D--------- ---------- ---------- ---------- H--------- ---------- ----------
50-4(2)-A5   D--------- ---------- ---------- ---------- ---------- ------C--- ----------
50-4(2)-A6   D--------- ---------- ---------- ---------- ------K--- ---------- ----------
50-4(2)-B3   D--------- W--------- ---------- ---------- ---------- ---------- ----------
50-4(2)-B5   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-B6   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-C2   ---------- ---------- ---------- ---------- ------K--- ---------- ----------
50-4(2)-C4   D--------- ---------- ---------- ---------- ---------- ---------- ----------
50-4(2)-C5   ---------- ---------- ---------- ---------- -----L---- -------L-- ----------   140


IIIB_Vif     KVGSLQYLAL AALIKPKQIK PPLPSVRKLT EDRWNKPQKT KGHRGSHTMN GH*
50-4(1)-A1   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-A6   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-B2   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-B3   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-B5   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-C2   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-C5   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-D1   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-D2   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-D3   ---------- ---------- ---------- ---------- ---------- ---
50-4(1)-D4   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-A1   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-A4   ---------- ---------- ---------- ---------- ---------- R--
50-4(2)-A5   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-A6   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-B3   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-B5   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-B6   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-C2   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-C4   ---------- ---------- ---------- ---------- ---------- ---
50-4(2)-C5   ---------- ---------- ---------- ---------- ---------- ---   140
```

Figure 5-8: Alignment of HIV$_{IIIB}$ *vif* alleles selected by rhA3F.

218

```
LAI_Vif        MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW
50-5(1)-A5     men------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-A6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-B1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-D2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-D3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-D4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-A1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-A2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-A4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-A5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------          70

LAI_Vif        GLHTGERDWH LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN
50-5(1)-A5     ---------- ---------- ---------- ---------- ------N--- ---------- ----------
50-5(1)-A6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-B1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-D2     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-D3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(1)-D4     ---------- ---------- -R-------- ---------- ---------- ---------- ----------
50-5(2)-A1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-A2     ---------- ---------- --K------- ---------- ---------- ---------- ----------
50-5(2)-A4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-A5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B3     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B5     D------E-- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
50-5(2)-C1     ---------- ---------- ---------- ---------- ---------- ---------- ----------          140

LAI_Vif        KVGSLQYLAL AALITPKKIK PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH*
50-5(1)-A5     ---------- ---------- ---------- ---------- ---------- ---
50-5(1)-A6     ---------- ---------- ---------- ---------- ---------- ---
50-5(1)-B1     ---------- ---------- ---------- ---------- ---------- ---
50-5(1)-B4     ---------- ---------- ---------- ---------- ---------- ---
50-5(1)-C1     ---------- ---------- ---------- ---------- ---------- E--
50-5(1)-D2     ---------- ---------M- ---------- ---------- ---------- ---
50-5(1)-D3     ---------- ---------- ---------- ---------- ---------- ---
50-5(1)-D4     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-A1     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-A2     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-A4     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-A5     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-B3     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-B4     ---------- ---------- ---------- ---------- ---------- E--
50-5(2)-B5     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-B6     ---------- ---------- ---------- ---------- ---------- ---
50-5(2)-C1     ---------- ---------- ---------- ---------- ---------- ---      192
```

Figure 5-9: Alignment of HIV$_{\text{LAI-GFP}}$ *vif* alleles selected by rhA3F.

```
IIIB_Vif       MENRWQVMIV WQVDRMRINT WKRLVKHHMY ISRKAKDWFY RHHYESTNPK ISSEVHIPLG DAKLVITTYW
51-3(1)-B1     -------I-- ---------- ---------- ---------- ---------- ---------- ----------
51-3(1)-B2     ---------- ---------- ---------- --K------- ---------- ---------- ----------
51-3(1)-B5     ---------- -------S-- ---------- ---------- ---------- ---------- ----------
51-3(1)-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(1)-C6     ---------- ---------- ---------- S--------- ---------- ---------- ----------
51-3(1)-D4     ---------- ---------- ---------- ---------- ---------- ---------- ------I---
51-3(2)-A1     ---------- ---------- ---------- ---------- ---------N ---------- ----------
51-3(2)-A4     ---------- ------N--- ---------- ---------- ---------- ---------- ----------
51-3(2)-A6     ---------- ---------- -------I- ---------- ---------- ---------- ----------
51-3(2)-B6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-C4     ---------- ---------- ---------- ---------- ---------- T------Q- ----------
51-3(2)-C5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-D2     ---------- ---------- ---------- ---------- ---------- ---------- ----------   70


IIIB_Vif       GLHTGERDWH LGQGVSIEWR KKRYSTQVDP DLADQLIHLH YFDCFSESAI RNTILGRIVS PRCEYQAGHN
51-3(1)-B1     D--------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(1)-B2     D--------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(1)-B5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(1)-B6     ---------- ---------- -------I-- ---------- ---------- ---------- ----------
51-3(1)-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(1)-D4     ---------- ---------- ---N----- ---------- ---------- ---------- ----------
51-3(2)-A1     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-A4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-A6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-B6     ---------- ---A----- ---------- ---------- ---------- ---------- ----------
51-3(2)-C4     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-C5     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-C6     ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-3(2)-D2     ---------- ---------- ---------- ---------- ---------- ---------- ----------   140


IIIB_Vif       KVGSLQYLAL AALIKPKQIK PPLPSVRKLT EDRWNKPQKT KGHRGSHTMN GH*
51-3(1)-B1     ---------- ---------- ---------- ---------- ---------- ---
51-3(1)-B2     ---------- ---------- ---------- ---------- ---------- ---
51-3(1)-B5     ---------- ---------- ---------- ---------- -------K-- ---
51-3(1)-B6     ---------- ---------- ---------- ---------- ---------- ---
51-3(1)-C6     ---------- ---------- ---------- ---------- ---------- ---
51-3(1)-D4     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-A1     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-A4     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-A6     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-B6     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-C4     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-C5     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-C6     ---------- ---------- ---------- ---------- ---------- ---
51-3(2)-D2     ---------- ---------- ---------- ---------- ---------- ---     192
```

Figure 5-10: Alignment of HIV$_{IIIB}$ *vif* alleles selected by rhA3G.

```
LAI_Vif      MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW
51-5(1)-A2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-A3   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-B1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C1   ---K------ ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C2   ---------- ----G----- ---------- ---------- ---------K ---------- ----------
51-5(1)-C3   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C6   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-D1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-D2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-D4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-A4   ---------- ---------- ---------- ---------- ---------- --------I- ----------
51-5(2)-A5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-B2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-B5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-C1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-C6   ---------- ---------- ---------- ---------- ---------- ---------- ----------        70


LAI_Vif      GLHTGERDWH LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN
51-5(1)-A2   ---------- ---------- ---------- ---------- ---------- ---------- -----R----
51-5(1)-A3   ---------- ---------- ---------- ---------- ---------- ---------- -----R----
51-5(1)-B1   ---------- ---------- ---------- ---------- ------V--- ---------- ----------
51-5(1)-C1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C2   ---------- -D-------- ---------- ---------- ------N--- ---------- ----------
51-5(1)-C3   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-C6   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-D1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-D2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(1)-D4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-A4   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-A5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-B2   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-B5   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-C1   ---------- ---------- ---------- ---------- ---------- ---------- ----------
51-5(2)-C6   ---------- ---------- ---------- ---------- ---------- ---------- ----------        140


LAI_Vif      KVGSLQYLAL AALITPKKIK PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH*
51-5(1)-A2   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-A3   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-B1   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-C1   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-C2   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-C3   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-C4   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-C6   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-D1   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-D2   ---------- ---------- ---------- ---------- ---------- ---
51-5(1)-D4   ---------- ---------- ---------- ---------- ---------- ---
51-5(2)-A4   ---------- ---------- ---------- ---------- ---------- ---
51-5(2)-A5   ---------- ---------- ---------- ---------- ---------- ---
51-5(2)-B2   ---------- ---------- ---------- ---------- ----R----- ---
51-5(2)-B5   ---------- ---------- ---------- ---------- ---------- ---
51-5(2)-C1   ---------- ---------- ---------- ---------- ---------- ---
51-5(2)-C6   ---------- ---------- ---------- ---------- ---------- ---         192
```
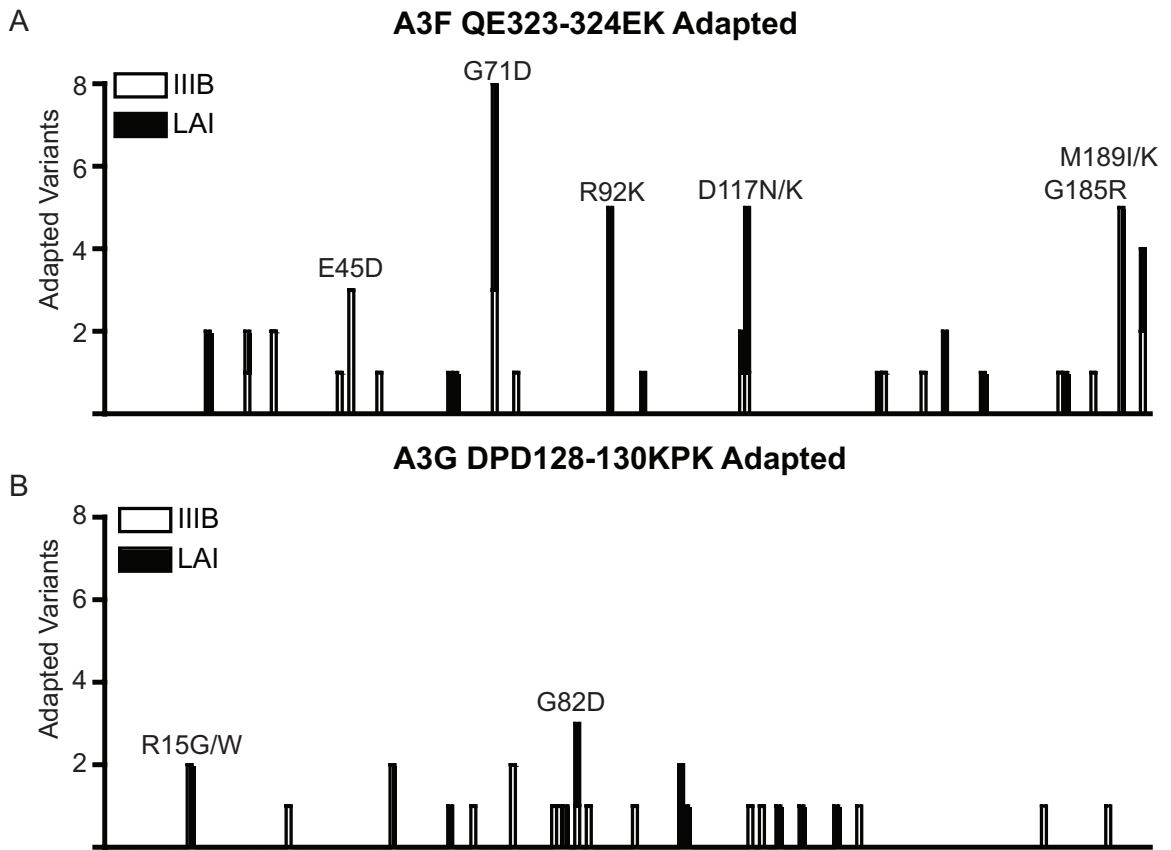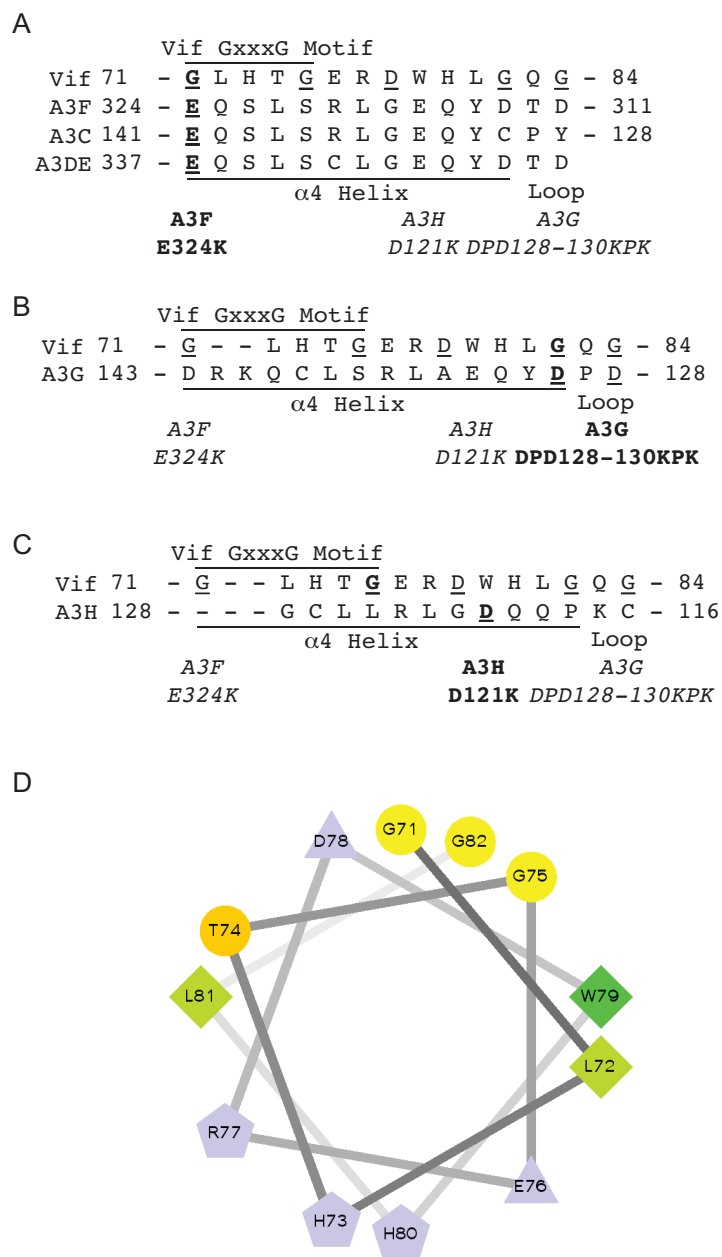
Figure 5-11: Alignment of HIV$_{LAI\text{-}GFP}$ *vif* alleles selected by rhA3G.

Figure 5-12: Mutational hotspots encountered in evolving *vif* for growth in the presence of A3F QE323-324EK or A3G DPD128-130KPK. (A) The frequency (y-axis) with which changes along the length of the *vif* open reading frame (x-axis) occurred in isolates evolved to grow in the presence of A3F QE323-324EK is shown. Hotspots are identified by the specific change encountered. **(B)** As in **(A)** but with viruses evolved to grow in the presence of A3G DPD128-130KPK. Minor populations assessed for A3F QE323-324EK-adapted viruses but not for A3G DPD128-130KPK-adapted viruses.
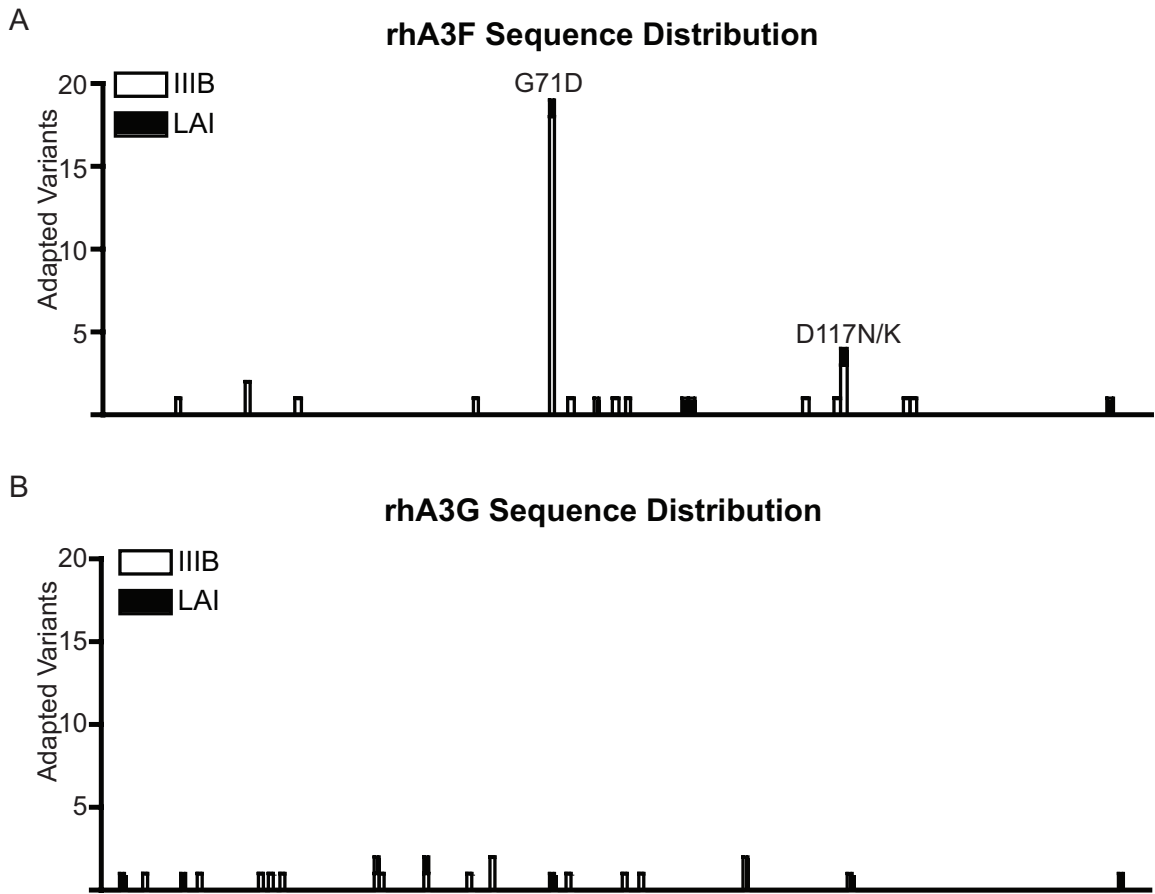
Figure 5-13: Mutational hotspots encountered in evolving *vif* for growth in the presence of rhA3F or rhA3G. (A) The frequency (y-axis) with which changes along the length of the *vif* open reading frame (x-axis) occurred in isolates evolved to grow in the presence of rhA3F is shown. Hotspots are identified by the specific change encountered. **(B)** As in **(A)** but with viruses evolve to grow in the presence of rhA3G. Minor populations not assessed.

A

```
          Vif GxxxG Motif
  Vif 71  - G L H T G E R D W H L G Q G - 84
  A3F 324 - E Q S L S R L G E Q Y D T D - 311
  A3C 141 - E Q S L S R L G E Q Y C P Y - 128
 A3DE 337 - E Q S L S C L G E Q Y D T D
                    α4 Helix        Loop
        A3F                  A3H         A3G
        E324K                D121K DPD128-130KPK
```

B

```
          Vif GxxxG Motif
  Vif 71  - G - - L H T G E R D W H L G Q G - 84
  A3G 143 - D R K Q C L S R L A E Q Y D P D - 128
                    α4 Helix            Loop
        A3F                  A3H         A3G
        E324K                D121K DPD128-130KPK
```

C

```
          Vif GxxxG Motif
  Vif 71  - G - - L H T G E R D W H L G Q G - 84
  A3H 128 - - - - G C L L R L G D Q Q P K C - 116
                    α4 Helix        Loop
        A3F                  A3H         A3G
        E324K                D121K DPD128-130KPK
```

D

Figure 5-14: A putative antiparallel helix in HIV Vif for interaction with the APOBEC3 α4 helix.

Figure 5-14: A putative antiparallel helix in HIV Vif for interaction with the APOBEC3 α4 helix. (A) A linear alignment of the putative antiparallel Vif helix with critical residues in A3F, A3D and A3C. Critical Vif-resistance (E324) and gain of function (G71) residues are bolded. **(B)** As in **(A)** but with A3G. **(C)** As in **(A)** but with a predicted compensatory mutation to counteract A3H D121K. Vif residues associated with a glycine-negatively charged amino acid pairing in this region are underlined throughout **(D)** Helical wheel representation of residues 71-82 in Vif derived from: http://rzlab.ucr.edu/scripts/wheel/wheel.cgi?sequence=ABCDEFGHIJLKMNOP&submit=Submit.
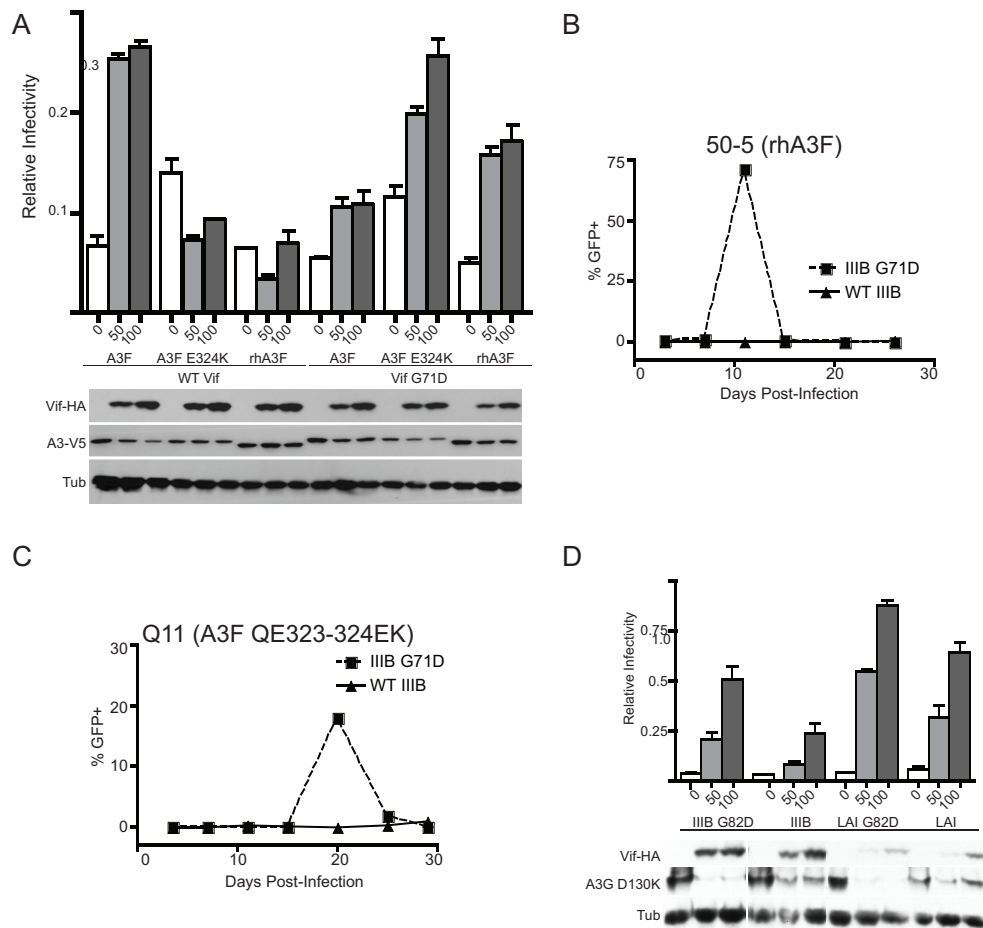
Figure 5-15: Vif G71D and Vif G82D neutralize Vif-resistant APOBEC3 proteins.
(A) A single-cycle experiment showing the infectivity and steady state levels of each APOBEC3 protein in the presence of either wildtype IIIB or G71D Vif. **(B)** Spreading infection data demonstrating the ability of HIV$_{IIIB}$ G71D to spread in rhA3F-expressing cells. **(C)** Spreading infection data demonstrating the more attenuated ability of HIV$_{IIIB}$ G71D to spread in cells expressing A3F QE323-324EK. **(D)** A single-cycle experiment showing the infectivity and recovery and steady state levels of A3G D130K in the presence of G82D versus wildtype IIIB and LAI Vifs.

226

Postscript

There are several aspects of this study that remain to be resolved before submission for publication, aside from higher MOI infections with molecular clone-derived mutant viruses and repetition of single-cycle experiments involving broader panels of selected mutants. First, we are expanding our analysis to test the efficacy of certain *vif* alleles against other Vif-resistant APOBEC3 proteins. Specifically, the equivalent of A3F E324 is conserved in A3C and A3D and renders at least the former of these Vif-resistant – the latter being indeterminate primarily because A3D expresses quite poorly and is thus difficult to assess (see **Supplementary Chapter 1**). We will therefore test the efficacy of G71D vs. the Vif-resistant variants of A3C (E141K) and A3D (E337K).

Similarly, the single amino acid determinant of Vif susceptibility associated with A3H falls between those associated with A3F and A3G. We would therefore predict that the interacting residue in Vif should also fall between the ends associated with A3F and A3G, or in other words, somewhere between G71 and G82. We will consequently be testing Vif G75D for the ability to neutralize A3H haplotype II D121K.

Finally, because the strongest result in this dataset is that associated with G71D and the suppression of Vif-resistant forms of A3F, there are two explorations of this topic that we are pursuing. One is a simple biochemical characterization of the sidechains at residue 71 capable of counteracting A3F E324K, since this goes to specificity of the direct interaction. The other, perhaps more critical, component is biophysical evidence for the ability of the predicted peptide to take on a helical structure and interact with A3F. Toward that end, we are collaborating with the Harki and Matsuo laboratories at the

University of Minnesota to create stapled peptides and test their association with the soluble form of the A3F C-terminal deaminase domain described in **Supplementary Chapter 1** by isothermal titration calorimetry. We anticipate that these complementary approaches will help to address any doubts that one might have about the existence of the proposed antiparallel helix and its direct interaction with the APOBEC3 α4 helix.

Chapter 6: Vif-independent Adaptation of HIV to Human APOBEC3 Proteins

Foreword

This is a working draft of a manuscript *in preparation* of the same name.

Albin JS and Harris RS. "Vif-independent Adaptation of HIV to Human

APOBEC3 Proteins." *in preparation.*

Whereas the adaptation experiments in **Chapter 5** were designed to define the putative

APOBEC3 interaction region of Vif corresponding to the Vif-interaction region located

in the APOBEC3 α4 helix, the purpose of the experiments described in this chapter was

to determine whether Vif-independent mechanisms of adaptation may be possible. This

was an important possibility to consider in light of our prior efforts demonstrating proof

of principle for just such an occurrence [(Haché *et al.*, 2008) and **Chapter 2**]. Such

mechanisms could be at play among some of the viruses described in **Chapter 5**, but to

unambiguously investigate the potential for Vif-independent adaptation of HIV to

APOBEC3 proteins, we felt it important to remove Vif from the equation. In **Chapter 6**,

we therefore utilize parent viruses containing large, presumably irreparable deletions in

their *vif* genes, as this bypasses any concerns about potential residual Vif activity that

might contribute to the overall adapted phenotype.

Two additional important modifications were made before commencing these

experiments to preemptively address problems that might arise based on our prior studies.

First, the original HIV$_{IIIB}$ virus contains an aberrant upstream open reading frame

adjacent to the primer binding site that interferes with viral production. When the virus

interrupts this upstream open reading frame by reverting to one of the nucleotides

presents in the overwhelming majority of patient isolates at position 200, increased viral

production titrates out encapsidated A3G in our fixed-expression cell lines (Haché *et al.*, 2009; Haché *et al.*, 2008). While interesting, this does not constitute a specific mutation relevant to the virus as it exists in patients. We therefore utilized only parent molecular clones in which the upstream open reading frame in question had been repaired. Doing so, however, would mean that all the resultant viruses would have to do to bypass A3G, based on our previous studies, would be to inactivate their *vpr* gene, which happens readily in culture. We therefore utilized SupT11-derived cell lines for our selections, since we had previously found that their levels of APOBEC3 expression restrict viruses with a repaired upstream open reading frame and a truncated *vpr* gene (**Chapter 2**).

Although we started from viral templates that lacked, to the best of our knowledge, unfit variations of the kind encountered previously, we did find in the course of these studies a partial p6 duplication in our $HIV_{LAI-GFP}$ virus that was sometimes deleted by adapted isolates and also rediscovered the importance of Nef, which is replaced by GFP in the same $HIV_{LAI-GFP}$ virus, for growth on certain cell types. For those reasons and because of their generally weaker Vif-independent phenotypes, we deemphasize the $HIV_{LAI-GFP}$ wing of these studies. The HIV_{IIIB} viruses used, however, do not suffer from any such defects of which we are aware.

Finally, we changed the selection protocol utilized in these studies to that described in **Chapter 5** to encourage more sequence diversification by giving the virus a chance to replicate in permissive cells while gradually increasing the APOBEC3-dependent selective pressure. This is closer in principle to the way drug resistance studies are typically performed than our former method of passaging viruses on fully nonpermissive cells from the start, which may be too stringent.

Summary

The HIV accessory protein virion infectivity factor (Vif) mediates the polyubiquitination and proteasomal degradation of the DNA cytosine deaminase APOBEC3G (A3G), thereby preventing the mutational activity of this host restriction factor from adversely affecting viral infectivity. Prior reports have demonstrated, however, that the Vifs derived from patient isolates show a wide range of anti-APOBEC3 activity, implying that the virus must have some way of averting APOBEC3 proteins in those situations where Vif is hypofunctional. To determine whether HIV can acquire Vif-independent adaptive mutations that enhance viral infectivity despite the presence of APOBEC3 proteins, we carried out genetic selection experiments designed to isolate Vif-deleted viruses that might utilize alternate routes of APOBEC3 evasion. Our results demonstrate that HIV can readily adapt to growth in the presence of A3G-expressing cells by acquiring mutations in the *env* gene. Surprisingly, we found that these mutations, despite the adaptive advantage conferred to the virus, do not appear to substantially alter levels of A3G encapsidation. We therefore propose that the HIV envelope may mediate an accessory, packaging-independent mechanism for counteracting A3G as a complement to the Vif-dependent reduction in A3G encapsidation.

Introduction

The human immunodeficiency virus type 1 (HIV) accessory protein virion infectivity factor (Vif) is required for the maintenance of viral infectivity *in vivo*, in primary T cells and in certain immortalized cell lines termed "nonpermissive" for Vif-deficient replication (Desrosiers *et al.*, 1998; Fisher *et al.*, 1987; Gabuzda *et al.*, 1992; Gabuzda *et al.*, 1994; Strebel *et al.*, 1987). Since the discovery of APOBEC3G (A3G) as a dominant-acting, Vif-suppressible restriction factor that can reconstitute the nonpermissive phenotype in otherwise permissive cells (Madani and Kabat, 1998; Sheehy *et al.*, 2002; Simon *et al.*, 1998), it has been generally assumed that the suppression of APOBEC3 restriction factors is the primary role of Vif. This assumption is consistent with the accumulation of G-to-A hypermutation during the infection of primates with viruses containing mutations in the Vif BC box or zinc coordinating motifs required for recruitment of the fully functional E3 ubiquitin ligase complex that mediates A3G polyubiquitination and proteasomal degradation (Schmitt *et al.*, 2010; Schmitt *et al.*, 2009).

Clinically-oriented studies on the potential roles of APOBEC3 proteins *in vivo*, however, suggest that the APOBEC3-Vif duality may be less straightforward than currently thought. A growing number of studies has yielded an increasingly unclear picture of whether or to what extent APOBEC3 proteins actively restrict HIV *in vivo* [(Amoedo *et al.*, 2011; An *et al.*, 2004; An *et al.*, 2009; Biasin *et al.*, 2007; Cagliani *et al.*, 2011; Cho *et al.*, 2006; De Maio *et al.*, 2011; Do *et al.*, 2005; Gandhi *et al.*, 2008; Jin *et al.*, 2005; Land *et al.*, 2008; Pace *et al.*, 2006; Piantadosi *et al.*, 2009; Pillai *et al.*, 2012; Ulenga *et al.*, 2008a; Ulenga *et al.*, 2008b; Valcke *et al.*, 2006; Vazquez-Perez *et*

al., 2009) and reviewed in (Albin *et al.*, 2010b; Sheehy and Erthal, 2012)]. Despite this, several analyses of patient isolates have indicated that various subtypes of Vif as well as variants within subtypes display a wide range of anti-APOBEC3 activity, with as many as 20% being defective for the neutralization of A3F, A3G or both (Albin *et al.*, 2010a; Binka *et al.*, 2012; Iwabu *et al.*, 2010; Simon *et al.*, 2005). Mundane explanations for this are possible. For example, some have found that the levels of A3F and A3H in primary cells are too low to effectively restrict HIV, thus obviating the need for a highly active Vif (Li *et al.*, 2010; Mulder *et al.*, 2010). Similarly, the virus may simply replicate in reservoirs where APOBEC3 levels are insufficient to greatly impact the overall health of the viral population.

That said, it is also possible that HIV may use accessory modes of APOBEC3 neutralization to survive when Vif is hypoactive. Such accessory mechanisms of evasion may hypothetically also be useful to the virus during the early production of virions from any given cell since Vif may not yet have effectively lowered APOBEC3 steady state levels at that time. These possibilities have gone largely unexplored to date. Although we have previously reported that increased viral production may be one mechanism by which HIV can "tolerate" A3G in the absence of Vif (Haché *et al.*, 2009; Haché *et al.*, 2008), this proof of principle is not applicable to the virus *in vivo* since the adapted viruses in that study merely repaired a defect in the parent virus that is not generally found in natural isolates.

For at least two other restriction factors, it is clear that alternative mechanisms of neutralization can be activated when need be. For example, it is now known that HIV and related lentiviruses use three different gene products to counteract the restriction factor

233

tetherin: HIV uses Vpu (Neil *et al.*, 2008; Van Damme *et al.*, 2008), certain simian immunodeficiency viruses (SIV) use Nef (Gupta *et al.*, 2009; Jia *et al.*, 2009; Sauter *et al.*, 2009; Zhang *et al.*, 2009) and HIV-2 uses gp41 (Le Tortorec and Neil, 2009). Intriguingly, when Nef-deficient SIV is passaged in primates, gp41 can reacquire anti-tetherin function (Serra-Moreno *et al.*, 2011), suggesting that ancient or perhaps simply masked alternative functions of lentiviral gene products may be biologically relevant. Similarly, the accessory proteins Vpr and Vpx can each act separately in different lineages to inactivate the myeloid cell restriction factor SAMHD1 (Lim *et al.*, 2012).

In both of these cases, the search for an alternative mechanism of restriction factor neutralization was propelled by the simple observation that the canonical viral defense (Vpu against tetherin, Vpx against SAMHD1) does not exist in closely related lentiviral species despite the fact that these viruses too must presumably encounter their host species' innate defenses. In the study of APOBEC3 proteins, however, such possibilities have gone largely unaddressed since *vif* is conserved in every known lentivirus with the exception of equine infectious anemia virus. If Vif is frequently hypofunctional or even nonfunctional, however, it is reasonable to hypothesize that HIV may have redundant mechanisms for preserving its genome from cytosine deaminases.

To determine whether HIV can acquire Vif-independent means of counteracting APOBEC3 restriction factors, we carried out genetic selection experiments aimed at isolating viruses with large, irreparable deletions in *vif* that nevertheless spread in cell lines expressing A3G. Consistent with our studies involving Vif-proficient viruses in the presence of Vif-resistant APOBEC3 proteins (**Chapter 5**), we find that Vif-deleted viruses can readily adapt to growth in the presence of A3G and that this phenotype

translates to the partial neutralization of A3F and Vif-resistant derivatives of A3F and A3G as well. Importantly, however, these viruses are not capable of spreading in naturally nonpermissive cells that express multiple APOBEC3 proteins, which may suggest that the selected adaptive mechanism is limited in its ability to counteract multiple APOBEC3 proteins or higher levels of a single APOBEC3 protein. On sequencing representative strongly resistant viruses, we noted a number of changes spread throughout the viral genome, which when recombined into molecular clone backgrounds revealed a novel anti-A3G activity for HIV Env. Despite their ability to spread in the presence of A3G, however, these viruses did not appear to substantially alter their levels of encapsidated A3G, suggesting that Env mediates an anti-APOBEC3 function capable of inactivating particle-associated A3G. Such a mechanism may represent an ancestral function of Env and may act in the present to neutralize any A3G that escapes Vif function. Understanding the mechanism by which Env functions with respect to the neutralization of APOBEC3 proteins will therefore be of great importance for our understanding of the interplay between lentiviruses and their host APOBEC3 proteins.

Results

Vif-deleted HIV can readily adapt to growth in the presence of A3G – To determine whether HIV can adapt to the presence of APOBEC3 proteins in the absence of Vif, we opted to utilize two viruses carrying large, presumably irreparable deletions in *vif* that would prevent *vif* repair of the kind previously observed when adapting viruses to APOBEC3 proteins [(Albin *et al.*, 2010a) and **Chapter 2**]. The first of these, $_{HIVLAI-GFP}$, was provided by Dr. Mario Stevenson (Miami University) and contains a 290 nucleotide in-frame deletion in *vif* as well as *gfp* in place of the *nef* open reading frame. The second carries a previously-described 230 nucleotide deletion and frameshift built into a variant of HIV$_{IIIB}$ from Dr. Michael Malim (King's College London) that has been engineered to ablate an aberrant upstream open reading frame that otherwise adversely affects virus production (Gibbs *et al.*, 1994; Haché *et al.*, 2009; Haché *et al.*, 2008).

A3G expression levels in the SupT11-derived A3G and A3G DPD128-130KPK cell lines used in this study are shown in **Figure 6-1A**; A3F expression levels are as shown in **Figure 5-1B**. In total, 88 cultures of Vif-deleted virus were subjected to selection in the combinations depicted in **Figure 6-1B** where each prong represents a 24-well plate with 22 selective cultures, an uninfected cell line control and a permissive infected cell line control. The selection strategy itself is shown schematically in **Figure 6-1C**. In brief, viruses were gradually passaged from permissive cultures to co-cultures made progressively more nonpermissive by increasing the proportion of cells expressing A3F or A3G, wherein approximately 8-11 days of viral replication intervened between each passage. At the end of each progression, a portion of Condition 3 cultures was passaged back to permissive conditions to undergo another round of selection as depicted

in Condition 4b, while another portion of Condition 3 cultures was passaged to 100% nonpermissive cells in Condition 4a for phenotypic assessment as indicated in **Materials and Methods** and, if indicated, saved for MOI-controlled confirmatory passages.

In total, cultures scored positive at Condition 4a were subjected to three additional passages from low MOI to confirm adapted phenotypes and purify their associated genotypes away from any residual unadapted viruses. The total number of adapted isolates present at the end of each round of selection as confirmed through fourth passage is depicted in **Figure 6-1D**. While we were readily able to isolate Vif-deleted viruses from cell lines expressing A3G, this was not the case for cell lines expressing A3F, echoing our prior selection studies in which viruses adapted to A3F universally repaired their *vif* genes while viruses adapted to A3G went though Vif-independent means to achieve growth [(Albin *et al.*, 2010a; Haché *et al.*, 2008) and see **Discussion**].

Adapted viruses can effectively neutralize A3G, A3F and Vif-resistant derivatives thereof – Comprehensive phenotypic scoring of adapted viruses is presented in **Tables 6-1** through **6-3**. Interestingly, and in contrast with our prior reports, most of these isolates that were well-adapted to growth in the presence of A3G displayed a weaker but appreciable adapted phenotype in cells expressing A3F as well as cells stably expressing Vif-resistant derivatives of A3F and A3G, QE323-324EK and DPD128-130KPK, respectively. Growth curves for four of these viruses are shown in **Figure 6-2**. Importantly, these viruses were not able to spread in naturally nonpermissive T cell lines such as CEM2n and they were generally delayed on our highest APOBEC3-expressing

lines relative to the lower-expressing clones, suggesting saturability of the putative Vif-independent adaptive mechanism (data not shown).

Viruses adapted to growth in the presence of A3F displayed similar but slightly more attenuated phenotypes, which may indicate that the mechanism underlying these adapted phenotypes is similar but, for reasons unknown, less efficiently selected by A3F. This could reflect generally higher levels of expression in our A3F lines relative to our A3G lines when using naturally nonpermissive cells as a benchmark for expression. Alternatively, if the putative Vif-independent mechanism of action were to function at the level of deaminase activity, it has been proposed that the deaminase-independent effects of A3F may be more robust than those of A3G [(Holmes *et al.*, 2007) and see **Supplementary Chapter 2**]. Given the low effective multiplicity of infection during the points at which an adapted variant first begins to spread within a mixed population, inefficacy against even weak deaminase-independent restriction mechanisms could account for some of the differential selection patterns between A3F and A3G. Regardless, given their clearer phenotypes, we opted to move forward only with viruses adapted to A3G. In particular, we chose HIV$_{IIIB}$-derived viruses because these were almost universally phenotypically stronger than HIV$_{LAI-GFP}$-derived viruses, perhaps due to *nef* deficiency in the latter. Moreover, we anticipated that the use of a single isogenic background would facilitate more valid comparisons among candidate adaptive mutations.

Mutational profiles of A3G-adapted viruses suggest exposure to A3G – To identify potential adaptive mutations in our viral isolates, we subjected four of the phenotypically

strongest HIV$_{IIIB}$-derived isolates (**Figure 6-2**) to clonal sequencing of their full genomes between the third and fourth confirmatory passages. As shown in **Figure 6-3A**, these isolates displayed extensive hypermutation in the 5'-GG-'3 dinucleotide context characteristic of A3G activity. Moreover, comparison of the levels of hypermutation found in different portions of the genome recapitulated the twin gradient pattern of hypermutation previously associated with A3G. Although we could not conclusively say from these data alone whether the adapted viruses were actively being hypermutated on an ongoing basis, it is clear that the isolates had been highly exposed to A3G, again validating the co-culture method employed. PCR analysis and sequencing also showed that the adapted viruses maintained their *vif* deletions, eliminating mundane explanations for enhanced growth in the presence of A3G such as contamination of the cultures with a wildtype virus (**Figure 6-3B**).

In analyzing the clonal sequences derived from this process, we noted a large number of apparently fixed mutations present in the genomes of the viruses chosen. In particular, 4/4 had inactivated their *vpr* genes, 4/4 had changes in *env*, 3/4 had changes in reverse transcriptase, 2/4 had changes in *gag* and 2/4 had altered their U3 promoter region. These mutations are depicted schematically in **Figure 6-3C** and in detail in **Tables 6-4** through **6-7**.

Changes in env enable enhanced spread by Vif-deficient HIV in the presence of A3G – To address the possibility that the U3 promoter mutations picked up by two viruses might alter viral gene expression and thereby enhance viral production to titrate out A3G as we have previously observed, we used a previously-characterized LTR-GFP reporter system

to determine the relative levels of gene expression derived from these viruses (Haché *et al.*, 2008). Cotransfection of these reporters containing the mutant U3 elements, however, failed to enhance GFP expression over the parent background (**Figure 6-4A**).

To characterize the potential contributions of other mutations in the viral genome, we made approximately 20 molecular clones comprising many mutations found together in a single viral gene as well as several combinations of these as they naturally occurred. On subjecting these molecular clones to spreading infection in A3G-expressing cells, however, we were surprised to note that only those viruses containing an altered *env* gene among the various genotypes tested were capable of enhanced spread in the presence of A3G (**Figure 6-4B**).

Selected env mutations do not substantially alter gross levels of encapsidated A3G – All known retroviral mechanisms of APOBEC3 evasion with the exception of that associated with equine infectious anemia virus involve the prevention of APOBEC3 encapsidation, whether by an active mechanism such as Vif-mediated degradation or by passively avoiding APOBEC3 encapsidation (Abudu *et al.*, 2006; Bogerd *et al.*, 2008; Derse *et al.*, 2007; Doehle *et al.*, 2006; Doehle *et al.*, 2005b; Löchelt *et al.*, 2005; Russell *et al.*, 2005; Zielonka *et al.*, 2009). To ask whether the envelope mutations selected by A3G act in a similar encapsidation-dependent manner, we infected SupT11-derived cell lines expressing the deaminase-deficient mutant A3G E259Q with molecular clone-derived viruses containing the *env* mutations gp120 A58V, gp120 P81T/gp41 T626M or gp120 M475I/gp41 L565Q as well as the parent HIV$_{IIIB}$ Vifdel virus and Vif-proficient HIV$_{IIIB}$. Because A3G E259Q does not effectively restrict HIV when stably expressed in T cell

lines, it can be used as a means for directly comparing the inherent A3G encapsidation potential among wildtype, Vif-deficient and other variants in an infectivity-independent fashion [(Browne *et al.*, 2009; Haché *et al.*, 2008; Miyagi *et al.*, 2007; Schumacher *et al.*, 2008) and see **Supplementary Chapter 2**]. At 9 days post-infection, we harvested viral supernatants from each culture, infected CEM-GFP reporter cells with a portion of each supernatant and then froze the remainder. Following fixation and flow cytometry of CEM-GFP cells, the frozen supernatants were thawed and purified through 0.45 um filters, and volumes of virus of equal infectivity as normalized by CEM-GFP were pelleted through 20% sucrose cushions to isolate virus particles in a manner analogous to that which we have previously reported (Haché *et al.*, 2008). As shown in **Figure 6-5A**, viruses carrying adaptive alleles of *env* encapsidate A3G at levels similar to that of the parent HIV$_{IIIB}$ *env/vif*(-) virus. Similar results were observed in parallel infections using two additional SupT11-derived cell lines expressing A3G E259Q as quantified in **Figure 6-5B**. Despite their enhanced infectivity in the presence of catalytically active A3G, then, the adapted alleles of *env* do not appear to substantially alter the gross levels of A3G encapsidated relative to the parental HIV$_{IIIB}$ *env/vif*(-) virus. It is conceivable that the minor, statistically insignificant decreases observed may be functionally relevant, but we cannot at present offer data in support of such a notion. Furthermore, the mobility of encapsidated A3G under denaturing conditions was typical of that observed in the presence of wildtype Env, suggesting that intravirion proteolysis of A3G is not involved in the Env-dependent mechanism described here despite a proteolytic component to the mechanism by which murine leukemia virus averts mouse APOBEC3 (Abudu *et al.*, 2006). Despite the lack of a significant reduction in the amount of A3G encapsidated, the

241

most common mechanism by which diverse retroviruses bypass APOBEC3 proteins (Abudu *et al.*, 2006; Derse *et al.*, 2007; Doehle *et al.*, 2006; Doehle *et al.*, 2005b; Löchelt *et al.*, 2005; Russell *et al.*, 2005), virion mislocalization to a space outside the viral core remains a mechanistic possibility, as does altered deaminase activity of the encapsidated enzyme.

Selected env mutations clusters to Layer 1 of the gp120 inner domain – To gain some structural insight into the potential functions of the mutations identified in our genetic selections, we mapped the mutations onto a known structure of gp120 containing the g41 interaction domain (Pancera *et al.*, 2010). Surprisingly, we found that the mutations selected cluster to loops on either side of helix 0 in Layer 1 of the gp120 inner domain in close proximity to residues of the gp41 interaction site. It is therefore possible that mutation of these conserved residues may alter the interaction of gp41 with gp120, although the mechanistic implications of this putative alteration are unclear.

Discussion

Here, we report the results of genetic selection experiments that have identified HIV Env as the mediator of a putative packaging-independent mechanism of A3G neutralization. This is unique among known instances of APOBEC3 evasion in retroviruses, as only the Vif-deficient lentivirus equine infectious anemia virus uses a similar, undefined packaging-independent anti-APOBEC3 mechanism (Bogerd *et al.*, 2008).

Previously, we have found that Vif-deficient HIV can tolerate the presence of A3G by enhancing its production to titrate out the encapsidated levels of A3G in fixed-expression T cell lines (Haché *et al.*, 2008). The mutations implicated in that study were a nucleotide change A200T/C and truncation of the *vpr* gene, the former of which is already present in most natural isolates and ablates an upstream open reading frame that interferes with viral translation and particle production (Haché *et al.*, 2009). The function of *vpr* truncation remains unknown.

In the studies described in **Chapter 6**, we sought to revisit the theme of Vif-independent adaptation to A3G to determine whether we might select changes that would further enhance our understanding of the restriction mechanism and shed light on the viability of targeting the APOBEC3-Vif interaction therapeutically. To do this, we designed our experiments to be somewhat different from those we had previously reported. In particular, we started with viruses that already had a pyrimidine at nucleotide 200, like the vast majority of natural isolates, and that had gross deletions in *vif* that would prevent *vif* restoration as a mechanism of adaptation (Albin *et al.*, 2010). Furthermore, we utilized SupT1-derived cell lines stably transfected with APOBEC3

proteins since we had previously determined that the levels of expression in these particular cell lines restrict T/C200, *vpr*(-) viruses and thus would force viruses to acquire mutations beyond the *vpr* truncations we had previously selected. Finally, we changed our selection protocol to a co-culture method that gradually increases the exposure of Vif-deficient viruses to APOBEC3 proteins, thereby more closely mimicking the methods typically used in drug resistance studies and allowing more sequence diversification prior to selection.

These modifications resulted in a somewhat distinct global picture of the study of HIV adaptation to APOBEC3 proteins compared with our previous findings. Rather than being rare events, we find that viruses can readily adapt the presence of APOBEC3 proteins under the new conditions described. It is not clear exactly what barriers must be bypassed by the virus *in vivo*, and thus we cannot speak to whether a gradual versus an immediate block such as that typical of our former method is more relevant to viral evolution, but it certainly appears that, for the *in vitro* study of alternative mechanisms of viral adaptation, a gradual approach will produce more candidate isolates for evaluation.

The phenotypes of the resultant isolates are quite strong when pitted against any single APOBEC3 protein and are clearly Vif-independent as evidenced by robust growth on Vif-resistant variants of A3F and A3G (**Figure 6-2**). Despite these robust phenotypes, we found that the isolates themselves were incapable of spreading in naturally nonpermissive CEM2n cells that express multiple APOBEC3 proteins. This may indicate saturability of the adaptive mechanism. That is, while the adaptive mechanism may successfully bypass one APOBEC3 protein, the expression levels of the full APOBEC3 repertoire in naturally nonpermissive cells may overwhelm the mechanism in question.

244

The inability of the selected viruses to spread in naturally nonpermissive cells, while reinforcing the importance of Vif and the potential therapeutic targeting of the APOBEC3-Vif interaction, does not negate the potential mechanistic insights that might be gained by further elucidating the underlying Vif-independent mechanism. Toward that end, we carried out extensive sequence and molecular clone analyses of the selected viruses that yielded the surprising result that the mutations conferring an adapted phenotype map to the *env* gene (**Figure 6-4B**). As described in **Chapter 2**, there are many points at which one might anticipate the potential for Vif-independent adaptation to APOBEC3 proteins, but Env, admittedly, seemed quite unlikely prior to these studies since the current understanding of the mechanisms of APOBEC3 restriction would indicate that APOBEC3 proteins should never come into contact with the extracellular portions of Env implicated by our data.

Our working model, based on the placement of the selected mutations on a known gp120 structure and the proximity of the mutated residues to the gp41 interaction site, is that these *env* mutations function to increase gp120 shedding and thereby enhance signal transduction that somehow alters A3G activity, perhaps at the deaminase level. This would be consistent with the existence of a deaminase inhibitor in T cells (Thielen *et al.*, 2007) and is the most plausible alternative given the fact that the extracellular domain of gp120 should theoretically never come into contact with A3G. Such an explanatory model might also account for the stronger phenotypes associated with the adapted isolates themselves relative to the molecular clone viruses. For example, if the mechanism were mediated through soluble gp120 accumulation, then the soluble gp120 present when infecting with adapted isolate supernatants but not with molecular clone-derived viruses

might be phenotypically relevant. This is, however, merely a hypothesis, and although we are actively pursuing said hypothesis, we have no direct evidence for such a phenomenon at present.

While we do not currently understand the mechanism at play, the finding that a gene product even more ancient than Vif can mediate anti-APOBEC3 function may imply that such activities were ancestral accessory functions of Env prior to the rise of Vif. In the present day, the subject of how much soluble gp120 accumulates in patients is a subject of some controversy, but a moderate position based on patient-derived samples would place the concentration in the range of hundreds of pg/mL (Cummins *et al.*, 2010; Klasse and Moore, 2004; Santosuosso *et al.*, 2009). If this is the case, it may be that Env actively functions *in vivo*, but not in a non-pressurized culture system in the absence of disruptive mutations, to alter A3G function and thereby assist Vif in counteracting A3G in lymphoid tissues. These suppositions remain, however, highly speculative, and additional work will be required to determine the precise mechanism by which Env variants may function to permit enhanced viral growth in the absence of Vif.

Materials and Methods

Plasmids – Coding sequences for the pcDNA3.1-derived plasmids used to make stably transfected A3F- or A3G-expressing cell lines are identical to those of NM_145298 and NM_021822, respectively. Mutations rendering either protein Vif-resistant or catalytically inactive were made my QuickChange site-directed mutagenesis (Stratagene) and have been described previously [(Albin *et al.*, 2010b) and **Supplementary Chapter 2**]. Full-length molecular clones of HIV$_{IIIB}$ and HIV$_{LAI-GFP}$ as well as Vif-deficient derivatives have been described previously (Albin *et al.*, 2010b). The parent HIV$_{IIIB}$ used is identical to that of EU541617 with the exception of an A200C nucleotide change introduced to disrupt an upstream open reading frame that would otherwise inhibit viral production (Haché *et al.*, 2008; Haché *et al.*, 2009). Derivatives containing mutations found in passaged isolates were made by subcloning fragments out of the PCR clones described below and into clean molecular clone backgrounds. LTR-GFP constructs were made by site-directed mutagenesis of the plasmids previously reported (Haché *et al.*, 2008). All constructs were verified by sequencing and by diagnostic restriction digest.

Cell lines – The APOBEC3-expressing cell lines used in these studies have been described previously. SupT11-derived cell lines expressing wildtype A3F or A3G were first reported in (Albin *et al.*, 2010a) and also appear in (Albin *et al.*, 2010b), where Vif-resistant derivatives were first reported. SupT11-derived cell lines expressing catalytically inactive A3G are described in **Supplementary Chapter 2**. Briefly, SupT11 cells were electroporated with linearized APOBEC3 expression plasmids, and clones were isolated by limiting dilution under G418 selection. APOBEC3 expression was

evaluated by using 250 μL of a lysis buffer consisting of 25 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM MgCl$_2$, 50 μM ZnCl$_2$, 10% glycerol, 1% Triton X-100, 50 μM MG132 and complete protease inhibitor (Roche) to lyse $5x10^6$ cells. Lysates were then mixed with a 5x sample buffer consisting of 62.5 mM Tris (pH 6.8), 20% glycerol, 2% sodium dodecyl sulfate, 5% β-mercaptoethanol and 0.05% bromophenol blue to a final 2x concentration of sample buffer. This mixture was boiled for 10 minutes prior to SDS-PAGE, and fractionated proteins were transferred to PVDF membranes and blocked in 4% milk dissolved in PBS 0.01% Tween. Blocked membranes were then probed with primary antibodies against A3F (#11474 from Dr. Michael Malim via the AIDS Research and Reference Reagent Program), A3G (#10201 from Dr. Jaisri Lingappa via the AIDS Research and Reference Reagent Program), or tubulin (Covance) followed by anti-rabbit or anti-mouse secondary antibodies (BioRad and Jackson) as appropriate. Prior to analysis of expression via each new antibody after the first, membranes were stripped at 50°C in a buffer consisting of 62.5 mM Tris (pH 6.8), 2% SDS and 100 mM β-mercaptoethanol.

Cell culture – Suspension T cell lines such as SupT11 and CEM2n were maintained in RPMI with 10% fetal bovine serum, penicillin/streptomycin and β-mercaptoethanol. 293T cells were cultured in DMEM with 10% fetal bovine serum and penicillin/streptomycin.

Adapted viruses – The selection procedure associated with the isolation of adapted viruses is described in the main text in and **Figure 6-1**. Candidate isolates were identified

248

qualitatively for the enhancement of one of any three characteristics relative to neighboring wells and to positive and negative infection control wells on each plate when cultured on 100% nonpermissive cells: CEM-GFP infectivity, syncytia formation or cell death. Isolates identified were then subjected to a total of three MOI-controlled confirmatory passages. Throughout, criteria for being labeled an adapted virus were as follows: i) Any virus identified in a particular round of infection must be blindly identified by the same qualitative criteria on 100% restrictive cells at the end of all subsequent rounds of infection and must appear adapted in all subsequent confirmatory passages. ii) During confirmatory passage, adapted viruses must rise to greater than 2% CEM-GFP infectivity within 21 days and must remain above 1% for at least one additional time point. iii) During the MOI-controlled second and third passages, adapted viruses must confirm on at least one of two cell lines expressing the APOBEC3 protein to which the isolate was adapted. iv) During the MOI-controlled fourth passage, adapted viruses must confirm on at least two of four cell lines expressing the APOBEC3 protein to which the isolate was adapted.

Spreading infections – Spreading infections were initiated by infecting 150,000 cells in 1 mL per well in 24-well plates with a volume of virus corresponding to a multiplicity of infection (MOI) between 0.01 and 0.05 as determined by CEM-GFP infectivity and indicated in the legend associated with a given experiment. CEM-GFP infectivity for molecular clone-derived viruses was determined by infecting 100 μL of 25,000 CEM-GFP cells per well in 96-well plates with increasing volumes of virus in a constant total volume of 250 μL; the volume of virus corresponding to a desired MOI was then

determined by linear regression. For passaged viruses where a single infectivity measure was available, a y-intercept of zero was used to carry out linear regression. Cultures were monitored periodically for the infectivity of their associated supernatants by infecting 25,000 CEM-GFP reporter cells in 100 μL RPMI with 150 μL of virus supernatant. Approximately 3-5 days later, CEM-GFP cells were fixed in 4% paraformaldehyde and analyzed by flow cytometry. Parent cultures were periodically split and fed to prevent cell overgrowth.

Viral genotyping – Viral genotypes were determined by analysis of sequences taken at or near the peak of viral spread during the third confirmatory passage. Half of infected CEM-GFP cultures in 96-well plates as used for standard monitoring (see above) were seeded into 500,000 uninfected CEM-GFP cells; several days later, genomic DNA was prepared from these seeded cultures using the PureGene genomic DNA isolation method. Genomic DNA was then treated with DpnI, and 550 ng of treated genomic DNA was used as the template for a touchdown PCR reaction amplifying fragments of approximately 1 kb each with the exception of the 2.2 kb *vif-vpr* region. PCR products were then gel-purified (Epoch or Fermentas gel extraction kits), cloned using CloneJet PCR cloning kits (Fermentas) and sequenced. Sequences were aligned and analyzed using Sequencer software (Gene Codes).

LTR-GFP reporter experiments – 293T cells were seeded in six-well plates at 250,000 cells per 2 mL per well. The following day, wells were transfected with 225 ng of plasmids in which the expression of GFP is dependent on the activity of a given

associated LTR (Haché *et al.*, 2008), 25 ng of a constitutively-expressing dsRed control plasmid and 250 ng of pcDNA3.1 vector plasmid to a total of 500 ng per well. Two days post-transfection, transfected cells were washed in PBS and analyzed by flow cytometry for total GFP fluorescence per total dsRed fluorescence.

A3G encapsidation – Spreading infections were initiated by infecting SupT11 cells stably transfected with the deaminase-deficient mutant A3G E259Q [*e.g.* (Schumacher *et al.*, 2008) and **Supplementary Chapter 2**] with wildtype or Vifdel HIV$_{IIIB}$ or with Vifdel derivatives carrying selected mutations in *env*. Nine days post-infection, supernatants were harvested from these cultures and used to infect CEM-GFP reporter cells prior to freezing the remaining supernatant. Following fixation of CEM-GFP and analysis by flow cytometry, frozen supernatants were thawed, filtered through 0.45 μm filters and normalized by volume according to their CEM-GFP infectivities. Virions were then centrifuged through 20% sucrose cushions, and pelleted viruses were resuspended in the 5x sample buffer described above. Western blotting for A3G (#10201 from Dr. Jaisri Lingappa via the AIDS Research and Reference Reagent Program) and for p24 (antibodies purified from hybridoma 183-H12-5C from the AIDS Research and Reference Reagent Program) was carried out as above, and images were quantified using ImageJ software.

Table 6-1: Phenotypic characterization of fourth passage, HIV$_{IIIB}$-derived, Vif-deficient A3G-adapted viruses.

Isolate	Growth on A3G	Growth on A3G DPD/KPK	Growth on A3F	Growth on A3F QE323-324EK	Round First Apparent
WT IIIB	++	-	++	-	na
IIIB Vifdel	-	-	-	-	na
C6-A1	++	+	+	+	3
C6-A2	++	+	+	+	3
C6-A3	++	+	+	+	3
C6-A4	++	+	+	+	3
C6-A5	++	+	+	+	3
C6-A6	+	+	+	+	3
C6-B2	+	+	+	+	3
C6-B3	+	+	+	+	3
C6-B4	+	+	+	+	3
C6-B6	+	+	-	+	3
C6-C1	++	+	+	+	3
C6-C2	++	+	+	+	3
C6-C4	+	+	-	-	3
C6-C5	++	++	+	+	3
C6-C6	+	+	+	+	3
C6-D1	++	+	+	+	3
C6-D4	++	++	+	+	2
C7-A2	++	+	-	+	2
C7-A3	++	+	+	+	2
C7-A4	++	+	+	+	2
C7-A5	+	+	-	-	3
C7-B1	+	+	nd	nd	2
C7-B2	++	+	+	+	2
C7-B3	+	+	-	+	3
C7-B4	++	+	+	+	2
C7-B5	+	+	-	+	2
C7-C1	+	+	-	-	2
C7-C2	+	+	+	+	2
C7-C3	++	+	+	+	2
C7-C5	+	+	+	+	3
C7-C6	++	+	-	-	2
C7-D1	+	+	+	+	3
C7-D2	++	+	+	+	2
C7-D3	+	+	-	+	3
C7-D4	++	++	+	+	2

Criteria

+ = Peaks at the last three time points (~3-4 weeks post-infection) or failure to peak on one or more cell lines used for each APOBEC3 expression condition
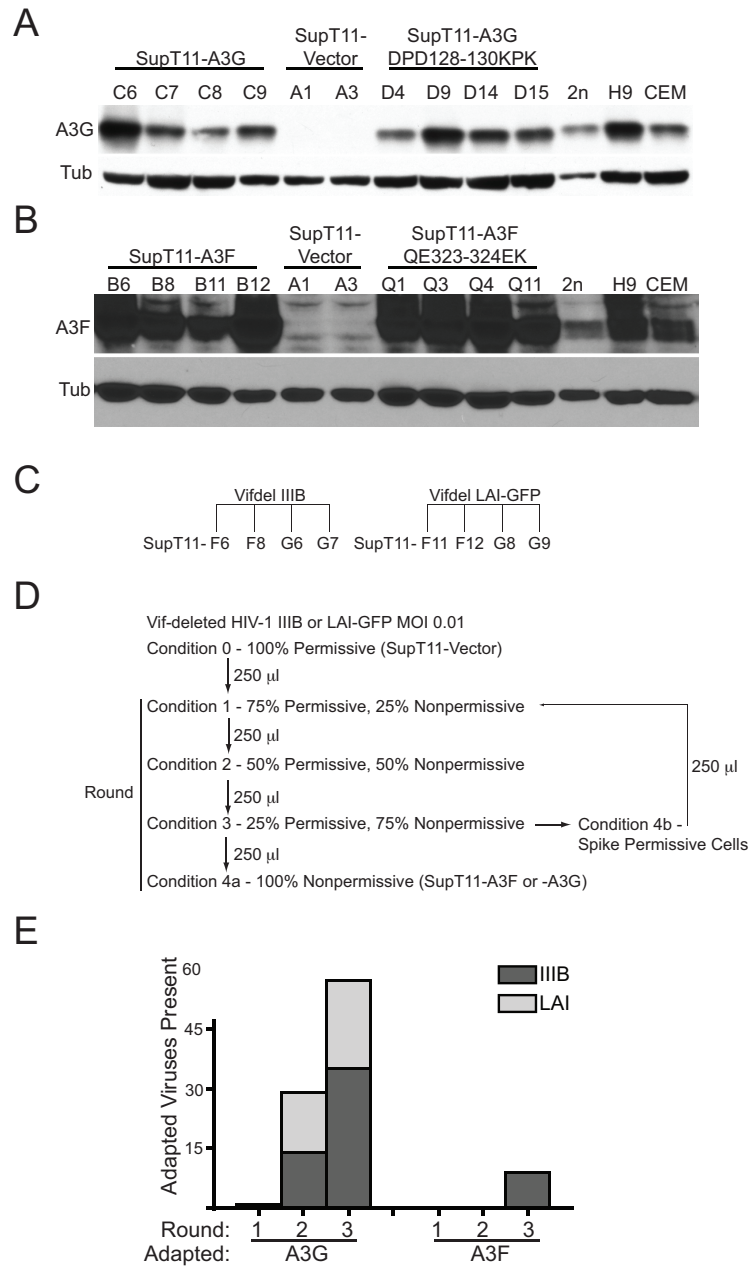
++ = Peaks at middle two time points (~1-2 weeks post-infection), peaks on at least ¾ A3G-expressing lines used; some viruses do not peak on the highest-expressing lines (C6 for A3G)

+++ = Peaks at first two time points (~1 week post-infection)

nd = Not determined, generally due to shortage of adapted virus to initiate infection.

na = Not applicable.

Note: "Peak" is defined not as the literal highest peak post-infection, but rather as the first peak of height at or comparable to the highest. Many isolates, likely due to being Vpr-deficient, will stay at high levels of CEM-GFP infectivity once they are up because they are less proficient at killing cells. Where there is disagreement (*e.g.* the peak on the highest level of A3G is more likely to be + versus ++ for the lowest levels of A3G), viruses are scored according to an overall impression of when viral growth is first occurring.

Table 6-2: Phenotypic characterization of fourth passage, HIV$_{\text{LAI-GFP}}$-derived, Vif-deficient A3G-adapted viruses.

Isolate	Growth on A3G	Growth on A3G DPD/KPK	Growth on A3F	Growth on A3F QE323-324EK	Round First Apparent
WT LAI-GFP	+++	-	++	-	na
LAI-GFP Vifdel	-	-	-	-	na
C8-A1	+	nd	nd	nd	3
C8-A2	+	+	-	-	3
C8-B3	+	-	+	-	2
C8-B5	+	-	-	-	2
C8-C1	+	-	nd	-	2
C8-C2	++	+	+	-	2
C8-D3	+	-	nd	-	2
C9-A2	++	+	+	-	2
C9-A3	++	-	nd	nd	2
C9-A4	++	+	-	-	2
C9-A5	+	-	-	-	2
C9-B1	+	-	+	-	2
C9-B2	+	+	nd	nd	2
C9-B3	++	+	+	nd	3
C9-B4	+	-	nd	nd	2
C9-B5	++	+	-	-	1
C9-C1	++	+	-	-	2
C9-C3	+	-	-	-	3
C9-C4	+	-	nd	nd	3
C9-C6	++	+	+	-	2
C9-D1	++	+	+	+	3
C9-D3	+	+	-	-	3

Criteria

+ = Peaks at the last three time points (~3-4 weeks post-infection) or failure to peak on one or more cell lines used for each APOBEC3 expression condition

++ = Peaks at middle two time points (~1-2 weeks post-infection), peaks on at least 3/4 A3G-expressing lines used; some viruses do not peak on the highest-expressing lines (C6 for A3G)

+++ = Peaks at first two time points (~1 week post-infection)

nd = Not determined, generally due to shortage of adapted virus to initiate infection.

na = Not applicable.

Note: "Peak" is defined not as the literal highest peak post-infection, but rather as the first peak of height at or comparable to the highest. Many isolates, likely due to being Vpr-deficient, will stay at high levels of CEM-GFP infectivity once they are up because they are less proficient at killing cells. Where there is disagreement (*e.g.* the peak on the

highest level of A3G is more likely to be + versus ++ for the lowest levels of A3G), viruses are scored according to an overall impression of when viral growth is first occurring.

Table 6-3: Phenotypic characterization of fourth passage, HIV$_{IIIB}$-derived, Vif-deficient A3F-adapted viruses.

Isolate	Growth on A3G	Growth on A3G DPD/KPK	Growth on A3F	Growth on A3F QE323-324EK	Round First Apparent
WT IIIB	++	-	++	-	na
IIIB Vifdel	-	-	-	-	na
B6-A6	+	+	+	+	3
B6-B2	+	+	+	+	3
B6-C3	+	+	+	+	3
B6-C6	+	+	+	+	3
B6-D3	+	+	+	+	3
B8-B6	+	+	+	+	3
B8-C1	+	+	+	+	3
B8-C2	+	+	+	+	3
B8-C6	+	+	+	+	3

Criteria

+ = Peaks at the last three time points (~3-4 weeks post-infection) or failure to peak on one or more cell lines used for each APOBEC3 expression condition

++ = Peaks at middle two time points (~1-2 weeks post-infection), peaks on at least 3/4 A3G-expressing lines used; some viruses do not peak on the highest-expressing lines (C6 for A3G)

+++ = Peaks at first two time points (~1 week post-infection)

nd = Not determined, generally due to shortage of adapted virus to initiate infection.

na = Not applicable.

Note: "Peak" is defined not as the literal highest peak post-infection, but rather as the first peak of height at or comparable to the highest. Many isolates, likely due to being Vpr-deficient, will stay at high levels of CEM-GFP infectivity once they are up because they are less proficient at killing cells. Where there is disagreement (*e.g.* the peak on the highest level of A3G is more likely to be + versus ++ for the lowest levels of A3G), viruses are scored according to an overall impression of when viral growth is first occurring.

Figure 6-1: Approach and successful isolation of A3G-adapted, Vif-deleted viruses.

Figure 6-1: Approach and successful isolation of A3G-adapted, Vif-deleted viruses.

(A) Western blots showing the A3G levels expressed in the cell lines used in this study. **(B)** Western blots showing the A3F levels expressed in the cell lines used in this study. **(C)** Schematic depicting the selection matrix used in this study in which each prong represents a 24-well plate with 22 selective cultures, one uninfected control culture and another control in which permissive cells were infected and the resultant viruses passaged continuously. **(D)** Schematic depicting the selection strategy employed as described in the main text. **(E)** Histograms demonstrating the number of adapted isolates present at the end of each round of selection as confirmed through a total of four confirmatory passages on 100% restrictive cells.
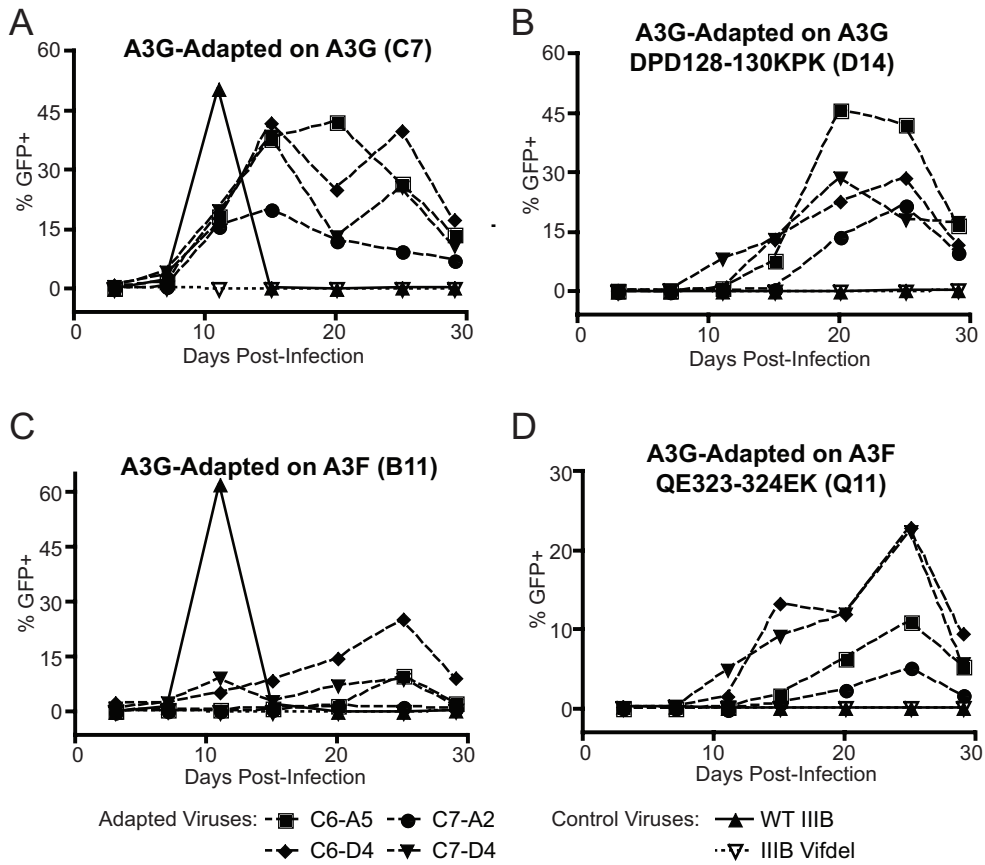
Figure 6-2: Phenotypes of adapted viruses. (A-D) Spreading infection curves show the fourth passage adapted phenotypes of the four isolates selected for sequence characterization.
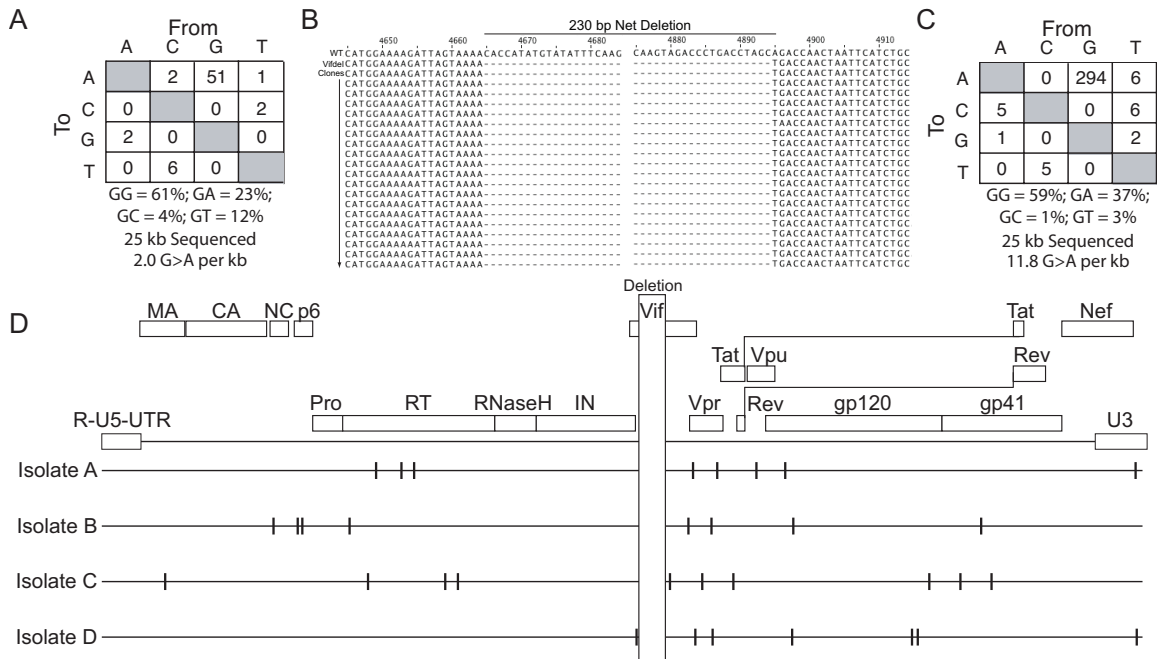
Figure 6-3: Sequence characterization of the selected viruses. (A) Hypermutation of fragments isolates from the 5' end of the genome. Contrast with levels found in fragments isolated from the 3' end of the genome **(C)** for overall mutation levels consistent with A3G action. **(B)** Direct sequencing across the virus demonstrates stability of the *vif* deletion employed. **(D)** Schematic depiction of the HIV genome showing the occurrence of the consensus mutations in the four isolates sequenced.

Table 6-4: Vif-deficient A3G-adapted A isolate C6-A5 consensus sequence[1].

Nucleotide Change[2]	Context[3]	Amino Acid Change[4]	Database[5]	Molecular Clones[6]
T2386G	CAT CC<u>G</u> GCA	RT P97P	n/a	n/a
G2610A	TTT A<u>A</u>A AAA	RT R172K	98% R; 1% K	R
T2724G	CTG T<u>G</u>G AGG	RT L210W	95% L; 1% W	L
C5135T	CCA <u>T</u>AG AGG	Vpr Q11X	n/a	n/a
ins5347.1A	<u>A</u>GG CGT TAC T	Vpr Frameshift	n/a	n/a
T5693G	GAA TA<u>G</u> AGG	Vpu Y29X	n/a	n/a
C5939T	GAT G<u>T</u>T AAA	gp120 A58V	98% A	A
delT8984	GGGACTT<u>T</u>CC	U3 NFκB II del8984	79% T; 21% del (varying size)	T8984

Table 6-5: Vif-deficient A3G-adapted B isolate C6-D4 consensus sequence.

Nucleotide Change	Context	Amino Acid Change	Database Distribution	Molecular Clones
G1503A	ACT ATT AAG	NC V390I	57% V; 21% I	V
AG1712–1713TC	CCT CAA GAG	p6 PE459–460PQ	61% A; 31% E	E[7]
C1750A	ACT CAC TCT	p6 P472H	70% P; 7% S	P
T2157C	AAA GCT AAA	RT V21A	99% V	V
G5097T	ACA TAG GAC	Vif E171X	n/a	n/a
C5300T	CAA TAA CTG	Vpr Q66X	n/a	n/a
C6007A	AAC ACA CAA	gp120 P81T	99% P	P
C7643T	CAC ATG ACC	gp41 T626M	93% M; 5% L; 1% T	IIIB/HXB2 T; LAI-GFP/NL4-3 M

1 Interpretation for HIV$_{LAI\text{-}GFP}$ is complicated by the fact that this virus carries a 12 aa duplication from the -1 to +11 positions of p6 (448-459). Sequencing of the p6 region from a greater number of adapted viruses than shown here demonstrated that some A3G-adapted HIV$_{LAI\text{-}GFP}$ isolates had deleted this duplication (data not shown).

Table 6-6: Vif-deficient A3G-adapted C isolate C7-A2 consensus sequence.

Nucleotide Change	Context	Amino Acid Change	Database Distribution	Molecular Clones
G562A	CGT A<u>A</u>A TCA	MA R76K	53% K; 40% R	R
G2318A	TTA <u>A</u>TA GAT	RT V75I	98% V	V
G2993A	GCA <u>A</u>AG CTA	RT E300K	97% E	E
G3088A	CAG AA<u>A</u> CAG	RT K331K	n/a	n/a
C4939A	GAA T<u>A</u>T GCT	Vif S118Y	n/a	n/a
del5215-5228	AT_ 12nt _TA	Vpr del5215-5228	n/a	n/a
T5486C	GTT TG<u>C</u> TTC	Tat C37C	n/a	n/a
G7191A	GAT AT<u>A</u> AGG	gp120 M475I	92% M; 6% I	M
T7460A	CAT C<u>A</u>G TTG	gp41 L565Q	61% L; 37% M	L
T7734C	AAG AA<u>C</u> GAA	gp41 N656N	n/a	n/a

Table 6-7: Vif-deficient A3G-adapted D isolate C7-D4 consensus sequence.

Nucleotide Change	Context	Amino Acid Change	Database Distribution	Molecular Clones
G4654A	AAA A<u>A</u>A TTA	Vif R23K	98% S	LAI-GFP/HXB2 S; IIIB/NL4-3 R
G5157A	GAA T<u>A</u>G ACA	Vpr W18X	n/a	n/a
A5312C	TTT <u>C</u>TC CAT	Vpr I70L	n/a	n/a
C6002T	GAC C<u>T</u>C AAC	gp120 P79L	99% P	P
A7042C	AAC <u>C</u>TG TGG	gp120 M426L	83% M; 6% R,L	M
A7091C	GGA C<u>C</u>A ATT	gp120 Q442P	30% N; 20% Q; 9-7-6% V-I-L	Q
C8988T	NFκB II G<u>C</u>TG NFκB I	U3 C8988T	79% C; 1% T; 1% A	C
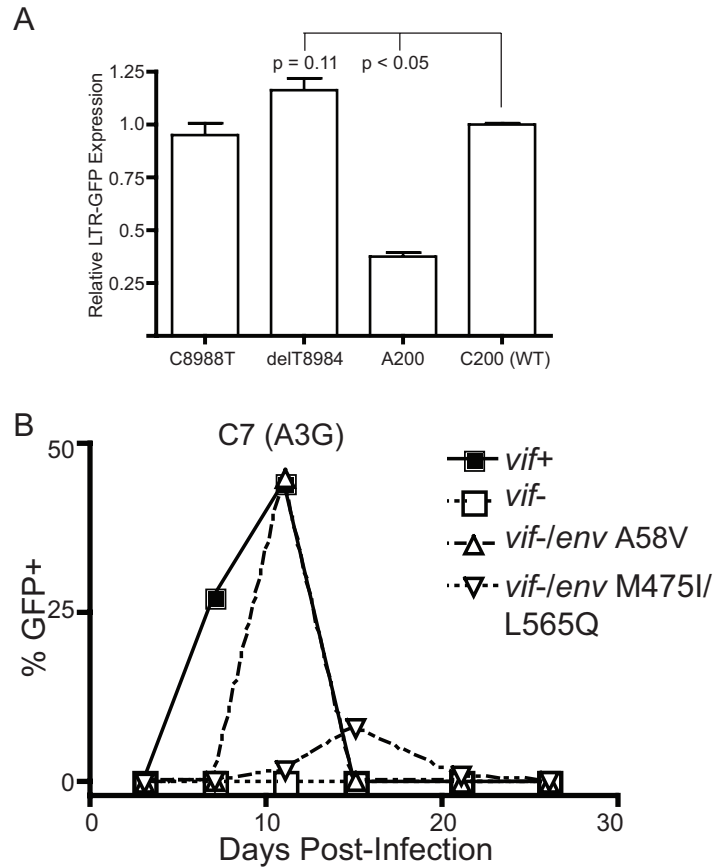
Figure 6-4: Mutations in *env* account for the adapted phenotype. (A) Histograms showing LTR-driven GFP expression relative to a dsRed control for two mutants identified in our selections, a known A200 hypomorph (Haché *et al.*, 2009; Haché *et al.*, 2008) and the parent C200 wildtype promoter. Data represent the mean and SEM of three independent experiments performed in duplicate. **(B)** Representative spreading infection curves showing that selected *env* mutations can overcome stably expressed A3G.
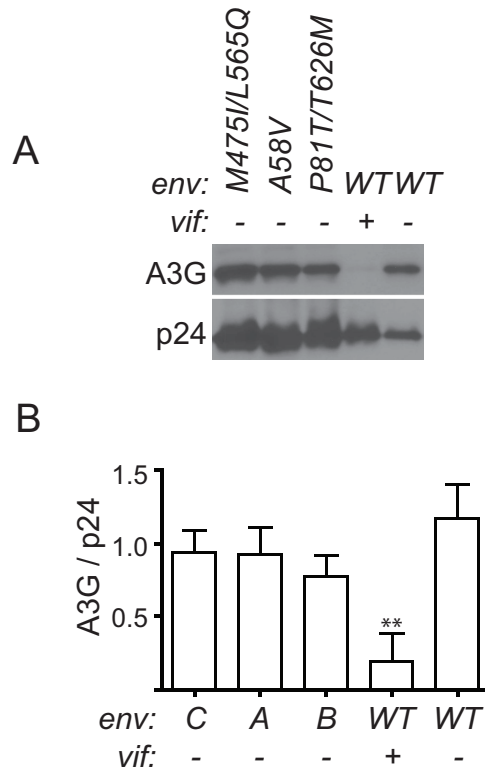
Figure 6-5: A3G encapsidation in selected *env* mutants. (A) Western blot showing the levels of A3G encapsidated by molecular clones of the indicated genotypes spreading at nine days post-infection on a SupT11 cell line stably expressing A3G E259Q. **(B)** Quantification of A3G encapsidation in adapted *env* backgrounds. A3G and p24 band intensities were quantified using ImageJ. Data represent the mean and standard deviation of A3G encapsidated per p24 for each virus on each of three, independent cell lines expressing A3G E259Q with the exception of *envB*, which was characterized on only two of these cell lines. One-way ANOVA with Dunnett's post test using the *envWT/vif(-)* virus as the control condition indicated a significant difference only between the control and *envWT/vif(+)* viruses. ** = $p < 0.01$
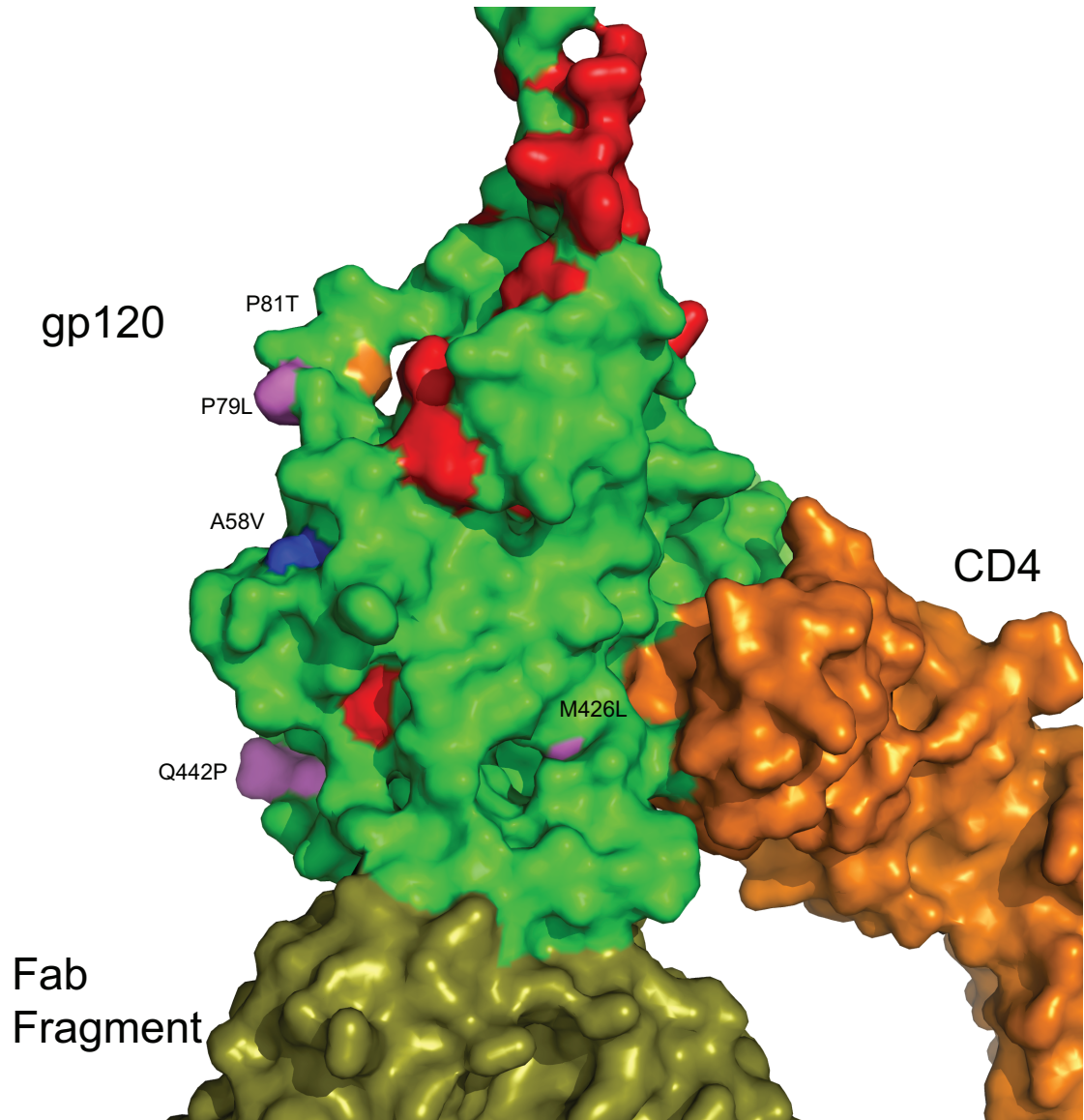
Figure 6-6: Selected *env* mutations cluster to loops on either side of helix 0 in layer 1 of the gp120 inner domain. Crystal structure 3JWO and summary of gp41-interactive residues (red) are derived from (Pancera *et al.*, 2010). Different colors correspond to distinct isolates. Not all residues changed are to be found in extant structures.

Postscript

There are several major unresolved issues that must be addressed in this study. First, we fundamentally do not know the mechanism by which Env counteracts A3G. Although an alteration to the amount of A3G encapsidated would be the most likely mechanism that one might predict, this appears not to be the case per our present data. One alternative possibility is that Env may alter the deaminase activity of particle-associated A3G, which is a plausible mechanistic basis that we are actively investigating. Another potential mechanism of Env function may be to alter the way in which the virus can spread by cell-to-cell contact, perhaps bypassing A3G in the process. To address this possibility, collaborators are presently investigating the cell-cell spread of the Env mutants described in this study. Finally, virion mislocalization remains a possibility since the localization of A3G outside of the viral core may render it unable to exert its antiviral effects.

Basic characterization of the *env* alleles selected is also ongoing. Specifically, selection for changes at highly conserved residues near the gp41 interaction site in gp120 implies the potential for an altered interaction between these two extracellular components. Although we do not know what the nature of such an alteration may be, it is possible that mutation of these conserved residues may weaken the interaction between gp41 and gp120, perhaps leading to increased levels of soluble gp120 in culture. Thus, we are investigating the levels of intact gp120:gp41 complexes among the selected mutants in comparison with the parent Env.

Finally, there are several experiments pertaining to the generalizability of the ability of the Env mutants selected to counteract A3G. All of the cell lines stably

transfected with A3G in this chapter are derivatives of SupT1. It will therefore be important to determine whether the same phenotype is displayed by some of our cell lines derived from different parents such as CEM-SS. Beyond that, it is clear at this point that the *env* mutant molecular clones are not phenotypically equivalent to the passaged isolates from which they derive, an observation that similarly holds true for the viruses described in **Chapter 5**. We do not presently understand this, but several possible explanations merit exploration. For example, it may be that accessory mutations in the natural isolates aid Env in neutralizing A3G; to address this, we will need to analyze molecular clones with combinations of multiple selected mutations built back in to clean molecular clone backgrounds. It may also be that a substance in the passaged isolates important for the neutralization of A3G accumulates over time. This could be, for example, soluble gp120 if there is a partial disruption to the interaction between gp120 and gp41. Such possibilities can be addressed by spreading infections using cultures pretreated with killed viral isolates, among other potential approaches.

In summary, many challenges remain in the elucidation of the mechanism by which the viruses described here function to bypass APOBEC3 proteins. We do not at present have data to support a mechanistic model that fully explains the phenotypes of our passaged isolates. Despite these challenges, the localization of APOBEC3-adaptive changes to an open reading frame other than *vif* is very important in that it suggests that we will need to consider the potential for accessory mechanisms of APOBEC3 evasion in implementing potential APOBEC3-related therapies.

Chapter 7: Conclusions and Discussion

The bulk of the data presented in this thesis describes a stepwise progression through the determination of important components required for the Vif-mediated neutralization of APOBEC3 proteins. In particular, the data may be conceived of as addressing five key questions: 1) Which APOBEC3 protein may best serve as a model for Vif-mediated neutralization? 2) Where in this APOBEC3 protein does Vif interact? 3) What are the downstream ubiquitination sites within this APOBEC3 protein required for its ultimate proteasomal degradation? 4) Where in Vif does this APOBEC3 protein interact? 5) If one could prevent Vif from facilitating the degradation of APOBEC3 proteins, how might the virus adapt? These questions and their respective answers per the work presented here are described visually in a schematic of the Vif-associated E3 ubiquitin ligase complex shown in **Figure 7A-B**. In viewing the data through this lens, the reader may appreciate how the various components of this dissertation unite in the central goal of addressing how one might go about the therapeutic utilization of APOBEC3 proteins.

Chapter 2: Long-term restriction by APOBEC3F selects human immunodeficiency virus type 1 variants with restored Vif function – Prior to the study described in Chapter 2, we had carried out experiments aimed at determining whether HIV might be able to adapt to growth in the presence of A3G without Vif, thus implicating potential alternative pathways around APOBEC3 proteins that might yield insight into the mechanism of restriction and into the therapeutic potential of APOBEC3 proteins (Haché *et al.*, 2008). These studies utilized a parent HIV$_{IIIB}$ containing tandem stop codons in *vif* to ablate

expression of the viral accessory protein, and despite extended passage in the presence of T cells stably expressing A3G, rare adapted variants of these viruses never altered these tandem stop codons, indicating that HIV might be capable of acquiring Vif-independent resistance to APOBEC3 proteins. In this specific case, adapted viruses acquired a noncoding nucleotide change A200T/C and a defective *vpr* open reading frame, the former of which enhances viral particle production to effective titrate out the levels of encapsidated A3G and allow the virus to survive at a population level while still being heavily exposed to A3G (Haché *et al.*, 2009).

One caveat to the major result of the aforementioned study, however, is that the viruses selected in these studies were not capable of growing in cells stably expressing A3F or in naturally nonpermissive CEM cells expressing multiple APOBEC3 proteins. We therefore carried out a companion series of experiments aimed at further adapting the A3G-selected viruses to also grow in the presence of A3F. These efforts yielded the surprising finding that, in contrast with our A3G studies, passage in the presence of A3F always resulted in strong selection for reversion of the tandem stop codons in *vif* to a variety of sense codons (**Table 2-2**). Thus, we were able to readily detect genetic interaction between A3F and Vif but not between A3G and Vif, suggesting that Vif is absolutely required for neutralization of A3F but not A3G. This may be taken to further imply that the A3F-Vif interaction may be a superior model for the study of the APOBEC3-Vif interaction, as would prove to be the case in the years following (**Figure 7-1**).

The exact nature of the reverted stop codons in *vif* turned out to be important for the interpretation of our selection studies. That is, the *vif* alleles selected by A3F proved

incapable of neutralizing A3G (**Figure 2-5**). A3F therefore exerts selective pressure for the maintenance of HIV Vif function distinct from that thought to be exerted by A3G. While it is not possible to extrapolate perfectly to the *in vivo* situation based on our culture models, this result implies that A3F is an important component of the selective pressures that maintain Vif function in the context of natural infection.

When we sought to define the nature of this differential Vif efficacy, we found that the identity of residue 26 in Vif was critical for the ability of the accessory protein to neutralize A3G. Despite the fact that all tested identities of both residues 26 and 27 in Vif were capable of neutralizing A3F, a lysine was required at position 26 to neutralize A3G (**Figure 2-6**). This was consistent with two directed mutagenesis studies of Vif residues critical for anti-APOBEC3 function that came out shortly before our manuscript was published (Chen *et al.*, 2009; Dang *et al.*, 2009). The mechanism by which this amino acid functions in the selective neutralization of A3G is not known. We observed no appreciable difference in the ability of Vif with K26 versus other residues at position 26 to coimmunoprecipitate A3G, and similar data were variable between the contemporaneous papers cited above. It is possible that this lysine could act as a site of autoubiquitination in Vif that could facilitate the degradation of bound APOBEC3 proteins as proposed previously (Dang *et al.*, 2008b), but based on our data presented in **Chapter 4** indicating that internal lysines in both A3F and A3G are the critical ubiquitin acceptor sites involved in APOBEC3 degradation, this seems unlikely. Combined with the existence of many other residues critical for the neutralization of one APOBEC3 protein but not another, then, this implies by default that K26 is somehow important for

the quality of the interaction between Vif and A3G but not other APOBEC3 proteins (**Table 1-2**).

In a broader context, **Chapter 2** is part of a recent debate in the literature on the restrictive power of A3F (Miyagi *et al.*, 2010; Mulder *et al.*, 2010). The crux of the matter is just how important A3F and other APOBEC3 proteins are relative to A3G, and the data in this chapter are among the most redeeming in favor of A3F. If a level of stably expressed A3F similar to that in naturally nonpermissive cells forces the virus to revert *vif* or fail to propagate, then the level of the threat from a viral perspective should be clear.

Arguing against A3F, however, are data suggesting that its expression levels in primary cells are so low as to be nonrestrictive (Mulder *et al.*, 2010). In this study, the authors utilized a Vif mutant thought to be selectively deficient for the neutralization of A3F. By comparing the unimpaired spread of the selectively deficient virus to that of a grossly deficient virus in primary cells, the authors concluded that the levels of A3F in primary cells must be too low to effectively restrict HIV. Certainly, this is possible and not at all an unimportant point to address, yet there are two important caveats to these data. First, there is no guarantee that the levels of A3F expressed in the primary cells used are indicative of the A3F levels present *in vivo*. While it is laudable to use primary cells, then, their ability to reflect *in vivo* conditions when placed in a culture dish is not guaranteed, nor is there any real evidence that there is one, true level of APOBEC3 expression since individual differences frequently cover a wide range of expression if one reviews some of the expression data discussed in **Table 1-1**. It also bears mentioning that parallel studies in which a selectively susceptible allele of *vif* is used to probe the efficacy

of A3G in primary cells have not been carried out. It is therefore not known to what extent these *ex vivo* models are indicative of restriction by any APOBEC3 protein, let alone A3F.

Overall, we fundamentally do not and cannot know "the" APOBEC3 repertoire of any one individual. Alleles vary. Expression levels vary. Cell types vary. The overall immune system varies. The nature of the virus infecting that individual varies, etc. All of these things will affect the expression and potential of APOBEC3 proteins *in vivo*, and thus to extrapolate the results of a handful of *ex vivo* primary cell experiments to an *in vivo* fact of life seems a false sense of physiologic security. In this case, "can happen" may be the best one can or should do until such time as a clinical trial becomes feasible subsequent to infrastructural "can happen" work.

Regardless, the reversions described in this chapter, repair of the *vif* open reading frame in a manner specific to the neutralization of A3F with a Vif that itself is nonfunctional against A3G, remain impressive in comparison with the broader scope of our efforts to adapt HIV to various APOBEC3 proteins. Three years and thousands of spreading infection cultures later, none of us in the Harris lab has yet seen anything quite that simple and clear despite numerous efforts to show a genetic interaction between A3G and *vif*. In following this lead by extending study of the APOBEC3-Vif interaction to A3F on top of A3G and other APOBEC3 proteins, we have gained important insights into shared and distinct features of these interactions.

Chapter 3: A single amino acid in human APOBEC3F alters susceptibility to HIV Vif – Based on the clarity of the genetic interaction observed in **Chapter 2**, we next sought to

determine where Vif interacts with A3F, as prior efforts in the field had focused only on critical Vif interaction residues in A3G or on loss of function mutants in Vif defective for the neutralization of one of more APOBEC3 proteins. To do this, we utilized the species specificity of the interaction between APOBEC3 proteins and Vif. That is, HIV Vif is capable of neutralizing human A3F but not rhesus macaque A3F (Virgen and Hatziioannou, 2007; Zennou and Bieniasz, 2006). By systematically substituting rhesus residues into the human protein where the two differ, we might thus be able to determine which determinants in A3F are critical for functional recognition by Vif, a strategy employed successfully with human A3G and African green monkey A3G (Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004).

Adhering to this method, we were able to determine that A3F E324 is a critical determinant of A3F interaction with Vif. In contrast with other reported single amino acid determinants of Vif susceptibility (D128 and D130 in A3G, D/E121 in A3H), residue 324 need not acquire a positive charge to render the protein Vif-resistant (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004; Zhen *et al.*, 2010). Rather, either a charge change at E324 or simply neutralizing the sidechain by mutation to alanine each renders A3F Vif-resistant (**Figure 3-4**). This demonstration of a positive requirement strongly implies that A3F E324 is a residue that directly interacts with Vif. This may also be the case for the other single amino acid determinants of Vif susceptibility, but it is difficult to assess direct interactions versus broader changes to the interaction surface caused by a mutation using the toolboxes employed by these authors.

Contemporaneous with the publication of our manuscript, another group using a different comparative approach identified A3F E289 as a similar determinant of Vif interaction, and mapping of these residues onto our model structure of the A3F C-terminus reveals that these residues are structural neighbors at the extreme C-terminal ends of the α3 and α4 helices (Smith and Pathak, 2010). This, combined with our inability to detect a quantitative difference in Vif binding between Vif-sensitive and Vif-resistant proteins, suggests that the residues identified in these respective efforts are part of a broader interaction surface (**Supplementary Chapter 1**). A recently reported structure of A3C with associated mutagenesis demonstrates the existence of just such a surface, perhaps explaining in part the variability in coimmunoprecipitation efficacy seen among different authors (Kitamura *et al.*, 2012). It may therefore be difficult to strictly correlate loss of anti-APOBEC3 function by mutation in either protein with loss of coimmunoprecipitation, although more sensitive, biophysical measures of association may shed light on this problem.

As demonstrated in **Chapter 5**, this straightforward finding yielded an important genetic tool that that we have since used to leverage insights well beyond this single amino acid. The basis for **Chapter 5**, however, was first reported in the manuscript associated with **Chapter 3**. That is, prior efforts aimed at mapping the determinants of Vif susceptibility in APOBEC3 proteins had emphasized the diversity of APOBEC3-Vif interactions (Bogerd *et al.*, 2004; Huthoff and Malim, 2007; Lavens *et al.*, 2010; Mangeat *et al.*, 2004; Russell *et al.*, 2009b; Schröfelbauer *et al.*, 2004; Xu *et al.*, 2004; Zhen *et al.*, 2010). We noted, however, that in comparing the phylogenetic equivalents of all of the single amino acid determinants of Vif susceptibility defined in A3G and A3H previously

that a pattern began to emerge. While all single residue changes were linearly distinct, mapping them onto our model structure of the A3F C-terminal deaminase domain revealed that these residues, in addition to all being negatively charged, exist on the surface of the α4 helix, which itself forms part of a broader negatively charged surface with α3. Thus, we were able to determine that there is a structural commonality to the Vif interaction region of APOBEC3 proteins underlying the seemingly divergent results emerging from various mapping studies. This raises the possibility that one might intervene therapeutically at one structural point to facilitate restriction by multiple APOBEC3 proteins, a point to which we return in our discussion of **Chapter 5**.

On the face of it, this seems logical. Vif is likely a flexible protein based on its apparent intrinsically disordered nature, but it would still be quite a lot for the approximately 90 amino acids that comprise the N-terminus that most studies implicate in APOBEC3 recognition to adapt to at least three different APOBEC3 interactions – A3F (and A3C and A3D, which typically behave as A3F does), A3G and A3H (Auclair *et al.*, 2007; Jäger *et al.*, 2011; Reingewertz *et al.*, 2010; Zhang *et al.*, 2011). Yet consensus in the field prior to this study and up to the present day would hold that this is exactly what happens. Whether the data described in **Chapter 5** will achieve moderation of this view only publication and time will tell, but those data strongly reinforce our suspicions about how Vif and APOBEC3 proteins interact based on a simple comparison of the critical residues found in different APOBEC3 proteins.

While an attractive model, however, it bears mentioning that there are many large gaps in our understanding of the APOBEC3-Vif interaction. On the APOBEC3 side of the interaction, the most prominent of these is what defines a Vif-susceptible deaminase

domain. Why are there no known instances of Vif recognition of a Z1-type deaminase domain? Why does Vif recognize one Z2, for example the A3F C-terminus, but not another such as the A3F N-terminus? An interesting and important genetic problem moving forward, then, will be to build on the α4 hook to expand out and define the broader elements that make up a Vif-susceptible deaminase domain – how to make the insensitive sensitive, and how to reliably predict novel Vif-resistance mutations. Such efforts will also facilitate future attempts at defining directly interacting residues in Vif corresponding to functionally altered APOBEC3 sites in a manner analogous to **Chapter 5**.

Chapter 4: Dispersed sites of HIV Vif-dependent polyubiquitination in the DNA deaminase APOBEC3F – Following Vif binding to an APOBEC3 protein, the next step in the degradation process is the polyubiquitination of downstream acceptor sites. This study sought to determine where such sites in A3F might occur with an eye toward testing the validity of a model proposed by a previous paper focusing on A3G (Iwatani *et al.*, 2009). In this model, binding of Vif to one cytosine deaminase domain orients the other for polyubiquitination at well-defined internal lysine acceptors, which is therapeutically significant in that it implicates a second potential target for intervention by preventing polyubiquitination of a given APOBEC3 protein rather than directly interrupting the interaction with Vif.

To address this model, we mutated all of the lysine residues in A3F to arginine and then systematically reverted single amino acids or several groups of amino acids. This yielded the surprising finding that A3F can be effectively polyubiquitinated in the

presence of Vif at a minimum of six distinct sites, and that these sites are dispersed throughout the length of the protein (**Figure 4-1** and **4-2**). Similarly, we demonstrated that the four lysines previously reported to be modified do not fully ablate A3G neutralization by Vif when changed to arginine, suggesting that alternative targets can be utilized (**Figure 4-1**).

To confirm the modification of these multiple residues in A3F and A3G, we collaborated with the Krogan lab at the University of California, San Francisco, which specializes in mass spectrometric methods. Using a technique in which all modified substrates are purified by binding to an anti-K-GG antibody to detect the isopeptide bonds typical of ubiquitination, they were able to confirm modification of many residues in A3F and in A3G distributed throughout both the N- and C-termini (**Figure 4-3**). The polyubiquitination of APOBEC3 proteins in the presence of Vif, then, appears to be a somewhat more structurally flexible process than previously thought.

Importantly, these mass spectrometry data represent the first direct evidence for the modification of any specific residue in A3F or A3G by Vif. Previous efforts have utilized a more traditional biochemical approach in which one cotransfects an APOBEC3 protein with or without Vif and a tagged ubiquitin, immunoprecipitates the APOBEC3 protein and then blots for ubiquitin, typically yielding a high molecular weight smear thought to indicate polyubiquitinated APOBEC3 substrate (Dang *et al.*, 2008b; Iwatani *et al.*, 2009; Shao *et al.*, 2010; Wang *et al.*, 2011b). This method leaves much to be desired, however, since one cannot know specifically whether the smear represents APOBEC3 proteins, associated coimmunoprecipitates (*e.g.* Vif) or some combination of the two. It is particularly concerning, given this sort of experimental design, that Vif has been shown

to autoubiquitinate itself [*e.g.* (Jäger *et al.*, 2011; Mehle *et al.*, 2004a)]. A more appropriate approach might be to blot for the APOBEC3 protein itself in this situation, but we have failed to detect a reliable APOBEC3-specific western blot signal of appropriate molecular weight under these conditions (data not shown). In any event, the mass spectrometry data are unambiguous and dovetail well with the genetically determined sites of modification.

Despite the apparently random linear distribution of targeted residues in our genetic and biophysical datasets, if one turns again to a structural understanding, the data begin to make sense. Surprisingly, the models of A3F and A3G available to us, which are based on a crystal structure of the A3G C-terminus, suggest that the lysine residues in these APOBEC3 proteins typically cluster to a common surface of the protein opposite the Vif interaction region [**Figure 4-4** and (Shandilya *et al.*, 2010)]. It is therefore possible that Vif, a highly basic protein, has evolved to recognize the negatively charged surface of APOBEC3 proteins located at and around the α4 helix, thereby positioning lysine-rich surfaces for downstream polyubiquitination (**Figure 4-6**). Such a model is consistent with the ability of the Vif-associated E3 ubiquitin ligase complex to target a wide range of residues since, in structural terms, these residues are of comparable accessibility. In the end, however, high resolution structures of the full APOBEC3-Vif-E3 ubiquitin ligase complex will be necessary to determine whether these indirect inferences are valid. Regardless of whether the binding model is correct, however, it is undeniable that a wide range of lysines can be effectively modified in A3F and A3G in the presence of Vif. Therapeutic targeting of this step in the degradation process will therefore likely prove intractable.

280

Chapter 5: Evidence for a conserved structural feature in HIV Vif that interacts with APOBEC3 proteins – Directed mutagenesis studies of the sort described in **Table 1-2** have suggested a variety of potential APOBEC3 interaction regions in Vif. If one sorts through these, several patterns emerge. First, regions that when mutated result in loss of function against A3F are often distinct from those that when mutated result in loss of function against A3G. This implies that there are aspects of the interaction between APOBEC3 proteins and Vif that are distinct from one APOBEC3 protein to another, and these selective deficiency regions are often presumed to represent the A3F- or A3G-specific binding regions, of Vif [*e.g.* (Smith *et al.*, 2009)].

In reality, none of these studies has reported direct evidence for the interaction of any particular residue in Vif with any particular residue in A3F or A3G. While it may be true that mutation of a particular region results in loss of function and perhaps loss of immunoprecipitation, these readouts do not necessarily mean that one has mutated a directly interacting region, as it is impossible to distinguish such a scenario from one in which the mutations made have caused a broader change in Vif structure. Based on these studies, then, unambiguous identification of the regions of Vif that directly interact with the presumptive Vif-interacting regions of the α4 helix in APOBEC3 proteins cannot be achieved.

Introducing the Vif-resistance mutations of the α4 helix into their respective APOBEC3 proteins yields an APOBEC3 variant capable of restricting HIV despite the presence of Vif. If the residues of α4 directly interact with a given residue in Vif, then, passage of wildtype HIV in their presence should eventually select for a compensatory

mutation in *vif* that would restore functional interaction, a concept illustrated in **Figure 5-1**. Such compensatory mutations represent superior evidence for direct interaction because they require a gain of function centered on a single point. There are, at least in theory, far fewer (perhaps as few as one) ways in which such a gain of function may occur, thereby limiting the myriad conceivable underlying mechanisms associated with loss of function mutations.

To test this, we chose Vif-resistant APOBEC3 proteins representative of changes at each end of the α4 helix, A3G DPD128-130KPK and A3F QE323-324EK. These choices carried an additional advantage for the identification of directly interacting residues in addition to be centered on a single point, namely their charge. That is, if a Vif-sensitive residue in an APOBEC3 protein becomes Vif-resistant by changing charge from negative to positive, then a directly interacting residue in Vif ought to be able to compensate for this by changing its own charge to something negative. Moreover, the choice of residues at each end of the α4 helix held out the possibility that one might identify through compensatory mutations in Vif not only the mutations themselves, but also a broader region important for interaction with α4 and associated structural insights.

By sequentially passaging Vif-proficient HIV in the presence of these Vif-resistant APOBEC3 proteins, we identified G71D and G82D as the most frequently occurring putative compensatory changes observed in viruses adapted to growth in the presence of A3F QE323-324EK and the related K324 rhesus macaque A3F or A3G DPD128-130KPK, respectively (**Figures 5-12** and **5-13**). The ability of G71D to functionally neutralize the E324K change was subsequently confirmed in single-cycle experiments and in spreading infections from clean molecular clone backgrounds (**Figure**

5-15). The G82D change was somewhat more difficult to assess, as this alone was only marginally selected for among A3G DPD128-130KPK-adapted viruses and did not display any appreciable effect against the selecting protein in our spreading infection or single-cycle systems (data not shown). We noted, however, that the G82D change itself was 11 amino acids away from G71D, which by linear alignment would place G82D opposite A3G D130 but not D128K. While the A3G D130K change is only partially Vif-resistant, we noted an improved effect of Vif G82D against A3G D130 over wildtype, suggesting that this compensatory change corresponds to A3G D130 (**Figures 5-14** and **5-15**).

As noted above, the 11 amino acids separating A3F E324 and D313 (the A3F equivalent of A3G D130) are equal to the 11 amino acids separating Vif G71 and G82. On closer inspection of this region, we further found that this 11 amino acid span contains a GxxxG helix-helix interaction motif and a total of four predicted linear pairs of a glycine with a negatively charged residue in the space of 11 amino acids, or five in 13 amino acids if one extends into predicted APOBEC3 loop space preceding the α4 helix [*e.g.* (Kleiger *et al.*, 2002; Lorieau *et al.*, 2010)]. Moreover, helical wheel modeling of this region shows that the four linear pairs contained within Vif 71-82 align to the same putative helical surface. Taken together, this implies that, rather than being separate interactions, the interaction of Vif with the APOBEC3 α4 helix is defined by a conserved, glycine-rich helix in Vif wherein the sterically accessible glycine sidechains may permit hydrogen bonding with inapposite negatively charged residues on the interacting helix [*e.g.* (Kleiger *et al.*, 2002; Lorieau *et al.*, 2010)].

283

Ideally, one would want both structural and genetic validation of the details of the APOBEC3-Vif interaction. The structural gold standard for the determination of such directly interacting regions would be a co-crystal composed minimally of Vif and an APOBEC3 protein. Unfortunately, such a structure remains technologically unfeasible. In the studies described here, however, we report positive results derived from the genetic gold standard – selection experiments for the identification of compensatory *vif* mutations. These have two major advantages over a crystal structure. First, the phenotypic validity of the defined interaction region is built into the experiment, minimizing questions over the relevance of the sometimes spurious interfaces associated with structural methods. Moreover, because our genetic data define a known helix-helix interaction motif, we can reasonably predict the structural basis of the APOBEC3-Vif interaction and therefore advance efforts to rationally intervene at this target long before a crystal will be available.

As of this writing, we are expanding on these observations by going in two directions. First, we hypothesize that the putative αG helix is a conserved Vif interaction motif that binds the APOBEC3 α4 helix. If this is so, then we would predict that a compensatory mutation capable of overcoming A3H D/E121K should also be found in this region and that, since the α4 Vif-resistance mutation in A3H falls between the equivalents in A3F and A3G, said neutralizing residue in the putative Vif interaction helix should also fall between G71 and G82. Second, for definitive proof of the existence of this helix-helix interaction in the absence of a co-crystal structure, we are pursuing biophysical evidence for the differential binding of a synthetic peptide corresponding to

the putative αG helix to Vif-resistant versus Vif-sensitive variants of A3F based on the A3F CTD variant with enhanced solubility described in **Supplementary Chapter 1**.

Aside from the specific region identified, our results are also important proof-of-concept for the notion of identifying directly interacting residues between Vif and APOBEC3 proteins by genetic selection. Based on the methods we have developed, one may extend selections to other recently-identified Vif-resistance mutations and thereby form an APOBEC3-Vif genetic interaction map over time in the absence of direct structural confirmation. Such efforts, in addition to potentially identifying new therapeutically targetable structural features, will be of great value in functionally assessing any co-crystal structures that the future may hold.

Stepping back from the details for a moment, it is important to emphasize that **Chapter 3** and **Chapter 5** unite in supporting the existence of conserved helices in each interaction partner facilitating the core of the APOBEC3-Vif interaction. This is not to say that the interaction of Vif with each APOBEC3 protein is identical, as both the large body of loss-of-function evidence and the observation that related but distinct compensatory amino acids overcome the Vif-resistance mutations of the α4 helix indicate that variations do exist. Rather, one might envision this putative core helix-helix interaction as the fulcrum around which Vif may teeter-totter in finding its final APOBEC3 binding configuration. This hypothesis is summarized in **Figure 7-2** in which we add to the data supporting compensatory mutations for Vif-resistant A3F and A3G the supposition that such a change will also exist to neutralize Vif-resistant A3H. The resultant "Trinity of APOBEC3-Vif interactions" conveys the overall message implicit in our data set against the background of the loss of function data summarized in **Table 1-2**:

285

while distinct from each other in certain aspects, all APOBEC3-Vif interactions share a core identity not unlike the Christian concept of three manifestations of one god, the trinity, from which the schematic derives.

Chapter 6: Vif-independent adaptation of HIV to human APOBEC3 proteins – The results of **Chapter 5** demonstrate that it is possible for HIV Vif to adapt to Vif-resistant APOBEC3 proteins. Adaptation through *vif* supports the direct interaction between APOBEC3 proteins and Vif and suggests that pharmacologic inhibition of this interaction may select for direct compensatory mutations at the interaction site. Such data are an absolute prerequisite in support of the therapeutic potential of the APOBEC3-Vif interaction per a maxim attributable to the influential retrovirologist John Coffin – to paraphrase, "If it doesn't select resistance, the drug's no good" (Haché *et al.*, 2008).

This is a generally valid way of considering drug targets wherein a small molecule inhibits a viral enzyme. If the small molecule doesn't select some kind of resistance mutation within the open reading frame encoding the targeted enzyme, then either the drug is ineffective or it does not work the way it is supposed to. Either way, it will be ineffective at inhibiting the virus, for obvious reasons in the former case and because the virus may simply acquire a second-site suppressor not subject to inhibition by the drug itself in the latter.

In considering therapeutic intervention via restriction factors, we might propose an addendum to the aforementioned rule of thumb that is true for both viral enzymes and restriction factors but somewhat easier to grasp in the latter case: "If it doesn't select *direct* resistance, the *target's* no good." In other words, we would expect an alteration to

the putative Vif binding site in APOBEC3 proteins to select direct compensatory mutations in *vif* since this is the only established interaction known to allow HIV to actively overcome APOBEC3 proteins. But what if HIV did not exclusively utilize Vif to bypass Vif-resistant APOBEC3 proteins? What if there were second site mutations capable of phenocopying Vif in the presence of APOBEC3 proteins? In addition to altering our mechanistic understanding of mechanisms by which APOBEC3 proteins restrict in a cell, such alternatives would likely be problematic for the use of APOBEC3 proteins as therapeutics, as the virus might simply acquire mutations at a second site not subject to one's intervention blocking the APOBEC3-Vif interaction and thus obviate the intervention itself.

It is clear from the results of **Chapter 5** that the selected mutations in *vif* cannot fully account for the phenotypes of the viral isolates described since many isolates carried either no obvious mutations in *vif* or mutations in *vif* that subsequently proved incapable of neutralizing the Vif-resistant APOBEC3 protein with which they were selected (data not shown). Although we did not test every *vif* mutation for efficacy against the Vif-resistant APOBEC3 protein that selected it, it is clear simply from the subset that we did that that most of these have marginal phenotypes at best, with the notable exception of mutations of the putative αG helix. These data strongly imply that HIV must be capable of acquiring alternative adaptive pathways around APOBEC3 proteins in much the same way that different lentiviruses utilize different genes to neutralize the restriction factor BST-2/tetherin [reviewed in (Douglas *et al.*, 2010)].

To unambiguously address the potential for Vif-independent mechanisms of adaptation to APOBEC3 proteins, we carried out selections of the kind described in

Chapter 5 using template viruses with gross, irreparable deletions in *vif* that would force the virus to act exclusively through second-site mutations in overcoming wildtype A3F or A3G, if such mutations should, in fact, exist. The result of these efforts is **Chapter 6** where we demonstrate that these deleted viruses can acquire Vif-independent strategies that effectively counteract APOBEC3 proteins, strategies that work, at least in part, through the viral *env* gene (**Figure 6-4**).

It is important to understand that ours is a field fundamentally based on a correlative observation – a Vif-suppressible protein is expressed in nonpermissive cells, and the ectopic expression that same protein can reconstitute the nonpermissive phenotype in otherwise permissive cells (Sheehy *et al.*, 2002). This is not to say that the correlation is invalid, but rather that it may be incomplete. In the cases of other restriction factors, the potential for different viral gene products counteracting the same restriction factor was immediately obvious due to the absence of the gene in other, closely related viruses (Gupta *et al.*, 2009; Jia *et al.*, 2009; Le Tortorec and Neil, 2009; Lim *et al.*, 2012; Neil *et al.*, 2008; Sauter *et al.*, 2009; Serra-Moreno *et al.*, 2011; Van Damme *et al.*, 2008; Zhang *et al.*, 2009), yet the strong conservation of *vif* in lentiviruses with the exception of equine infectious anemia virus has led to the narrow assumption that Vif function is the sole mechanism by which Vif-encoding viruses bypass APOBEC3 proteins. Looking at the data, however, one finds that it is not so simple. The range of anti-APOBEC3 potential among diverse Vifs is quite large, strongly implying that the virus may not always be able to depend on Vif to neutralize A3G and other APOBEC3 restriction factors (Albin *et al.*, 2010a; Binka *et al.*, 2012; Iwabu *et al.*, 2010; Simon *et al.*, 2005). If

this is the case, then how do these viruses carrying defective (from an APOBEC3 perspective) *vif* alleles survive *in vivo*?

Mundane explanations for such phenomena abound. The levels of APOBEC3 proteins expressed *in vivo* may be inadequate to fully restrict [*e.g.* (Mulder *et al.*, 2010)]. Alternatively, the virus may depend in part on low APOBEC3 expression reservoirs, or the defective alleles in question may skew our collective view by being disproportionately derived from these hypothetical reservoirs. It is also possible that HIV could acquire simple fitness mutations to outrun the levels of APOBEC3 expression it encounters *in vivo* in much the same way that we've previously shown it can outrun A3G overexpressed in cell lines (Haché *et al.*, 2008). All of these are possible, but it is also possible that HIV and other lentiviruses may have, at some point in their evolutionary history (and perhaps still), used alternate viral gene products to replace Vif function.

Our data supporting Env as an alternative anti-APOBEC3 factor, it may be fair to say, are quite literally the last mutations we would have expected to find in our selections. Since the extracellular portions of envelope never coexist with APOBEC3 proteins as we know them, it makes little sense to those of us conditioned to think of direct binding interactions and complexes to imagine that envelope could be involved in APOBEC3 neutralization. In reality, though, this is not as far-fetched as it may at first seem. There is an entire literature on functions of soluble gp120, and combining that with the known existence of an inhibitor of deaminases in T cells (Thielen *et al.*, 2007) and the apparent lack of a substantive effect on A3G encapsidation levels (**Chapter 6**), it is not a long leap to hypothesize signal transduction and subsequent inactivation of A3G as a mechanism of action. The trouble, of course, is proving it, and as of this writing, we are

unable to provide evidence speaking to a positive mechanism of action for Env in the neutralization of APOBEC3 proteins. It is possible that the very slight, statistically insignificant trend toward lower levels of A3G encapsidation might be functionally relevant, but it would be quite difficult to argue that such differences account for the phenotypes observed. Nevertheless, it is clear that a functional effect of Env, by whatever mechanism, does exist, and thus insight into said mechanism of Env neutralization of APOBEC3 proteins, as it becomes available, is likely to prove an important advance in our understanding of the APOBEC3 restriction mechanism.

Final Summary – In closing, this dissertation describes a systematic approach to investigation of the interactions of human APOBEC proteins with HIV Vif. By breaking these interactions down into digestible individual components, we have added to the field several important insights. One is an assessment of the functionally targetable features in the interaction. Namely, the APOBEC3-Vif interaction itself is likely a suitable target, while downstream polyubiquitination events are not. A second is the finding that specific viral proteins other than Vif, Env in this case, can compensate for the absence of the viral accessory protein when the virus has no other option. Unfortunately, we do not as of this writing understand the mechanism by which this occurs, but it will be important moving forward to understand how envelope and potentially other viral genes might effectively adapt to counteract APOBEC3 proteins.

In favor of continuing down APOBEC3-oriented therapeutic routes, however, these exotic alternative mechanisms of evasion appear insufficient to fully complement Vif function in naturally nonpermissive cells (data not shown). The known alternatives,

while of great mechanistic interest, are therefore unlikely to change the overall potential of APOBEC3 proteins as novel antiretroviral therapeutics. That said, we are also actively attempting to evolve Vif-deficient viruses that can neutralize the APOBEC3 repertoire present in naturally nonpermissive cells; if such viruses can be evolved, they may undercut the viability of the APOBEC3-Vif interaction as a therapeutic target.

For the moment, however, it appears that Vif is required to neutralize APOBEC3 proteins in the settings in which they normally occur – groups of several expressed in naturally nonpermissive cells. Importantly, we have identified common helix-helix interaction regions in APOBEC3 proteins and in Vif that appear to form the core of a common interaction that stands in stark contrast with the conventional wisdom in the field that Vif binds APOBEC3 proteins by several means. This common interaction region, while not yet fully validated, represents a highly attractive drug target for two reasons. One is that a single intervention may theoretically facilitate the antiviral activities of multiple APOBEC3 proteins. The second is more particular to the nature of the interaction. Because helix-helix interactions are a common means by which proteins contact each other, medicinal chemists have formed a wide variety of strategies for the creation of peptidomimetic small molecule scaffolds that can be tailored to the inhibition of specific protruding sidechains on a given target protein [reviewed in (Dewal and Firestine, 2011)]. If the Trinity Hypothesis is correct, then, it may form the basis for the rational design of APOBEC3-based therapeutics aimed at blocking the α4-αG interaction.
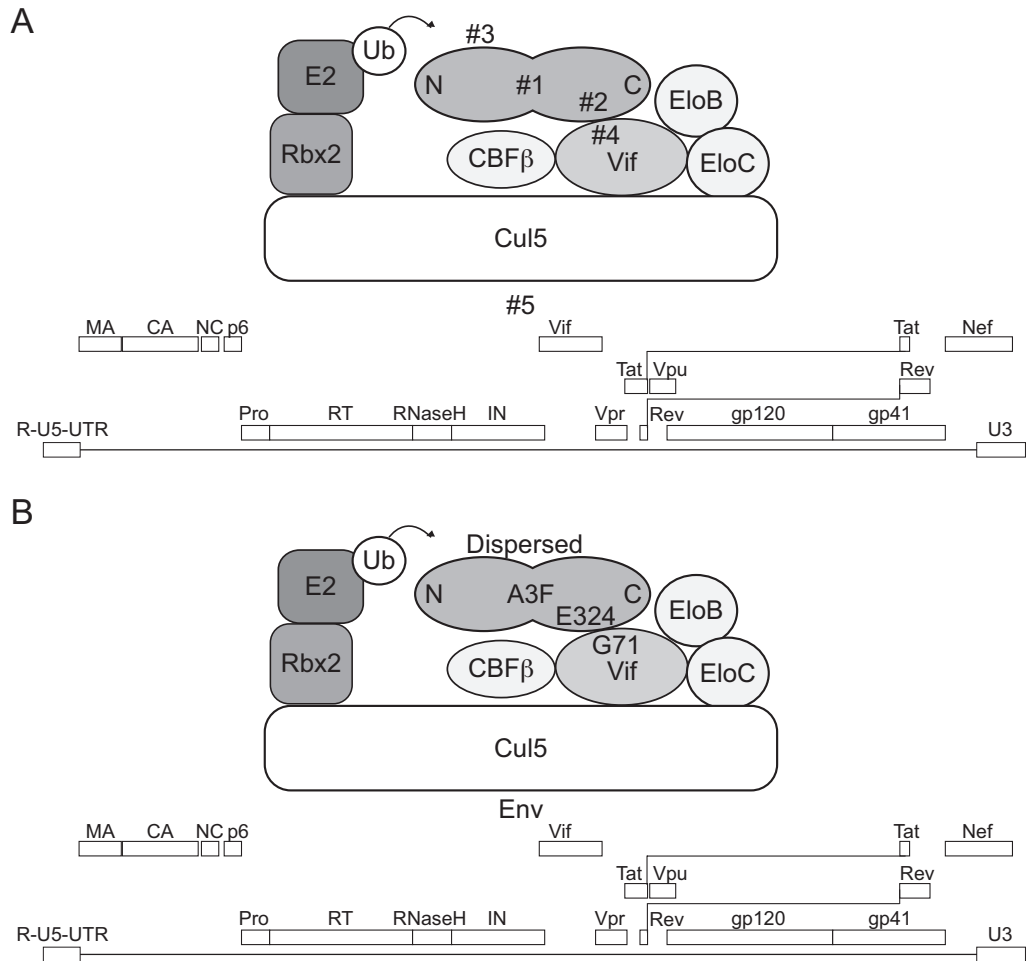
Figure 7-1: Schematic summary of the important findings reported in this dissertation. (A) A schematic depiction of the Vif-associated E3 ubiquitin ligase complex containing question marks associated with different components of the interaction. 1 = Suitable model APOBEC3 protein; 2 = Site of Vif interaction in that APOBEC3 protein; 3 = Downstream ubiquitin acceptor sites in that APOBEC3 protein; 4 = Vif residue interacting with the α4 helix of APOBEC3 proteins; 5 = Secondary pathways of HIV adaptation to APOBEC3 proteins. **(B)** Answers to the questions posed in **(A)** as reported in this dissertation.
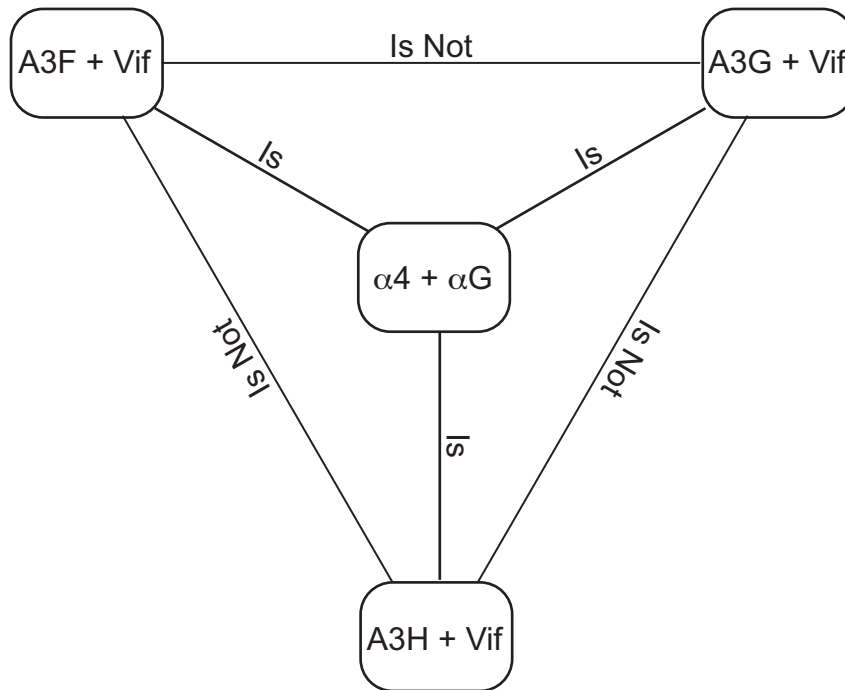
Figure 7-2: The Trinity Hypothesis. A schematic summary suggesting a common feature of the interaction between Vif and the α4 helix of Vif-sensitive APOBEC3 proteins.

Supplementary Chapter 1: Additional Mutagenesis of the A3F C-terminal

Deaminase Domain

Foreword

In Chapters 3 and 4 of this thesis, a series of mutagenesis studies establishing the Vif interaction site in A3F as well as the potential sites of subsequent Vif-mediated polyubiquitination were described. This supplementary chapter reports several additional mutagenesis studies of the A3F C-terminus performed either in the course of the work described in the main chapters or as functional validation of a series of mutants that has recently resulted in the successful crystallization of the A3F C-terminus.

Conservation of E324 in related Z2 domains – APOBEC3 subfamily cytosine deaminase domains cluster into three phylogenetic groups, termed Z1, Z2 and Z3 (LaRue *et al.*, 2009). While there is no known instance of Vif recognition of a Z1 domain, certain Z2 domains and the lone human Z3 domain are recognized by HIV Vif. The A3G N-terminal Vif recognition region is also part of a Z2 domain, while the A3G C-terminus is a Z1 type. In the case of A3F, both the N- and C-terminal deaminases domains are Z2 types despite the fact that Vif functionally recognizes only the C-terminal A3F Z2. Thus, A3F is similar to both the single-domain A3C and the double domain A3D in being solely comprised of Z2 domains.

As shown in **Figure SC1-1A**, E324 is conserved in A3C and A3D. We therefore hypothesized that mutation of the corresponding residue in these proteins would similarly stabilize them in the presence of Vif. Consistent with this hypothesis, the analogous mutation E141K in the single-domain A3C substantially enhances its stability in the presence of HIV Vif (**Figure SC1-1B**). Proper assessment of A3D was not possible in this experiment due to inadequate expression of the parent wildtype protein for comparison, but the fact that both Vif and A3D E337K were both substantially better expressed together than the wildtype combinations implies that the analogous change in A3D also stabilizes the related A3D C-terminal Z3 deaminase domain (data not shown). The structurally analogous A3G mutant D143K was not assessed, but we did note that changing the linearly analogous K141 to a glutamate failed to sensitize the otherwise Vif-resistant variant A3G D128K (data not shown). Thus, Vif recognition of the structural anchor point defined by A3F E324 is conserved in the related A3C Z2 domain and perhaps also in the A3D C-terminal deaminase domain, a matter currently under further

investigation as an extension of the data presented in **Chapter 5**. Interaction with the A3G N-terminal Z2 deaminase domain, however, appears at least partially distinct.

Additional mutagenesis of the A3F α3 and α4 helices – The observation that most known single amino acid determinants of Vif susceptibility map to a structurally conserved helix (**Chapter 3**) combined with the lack of a notable decrease in coimmunoprecipitation of the functionally Vif-resistant A3F E324K variant by Vif implied that a broader surface, perhaps involving the α4 helix as further elucidated in **Chapter 5**, might be recognized by Vif. To test this hypothesis, we mutated the polar surface residues of the A3F α3 and α4 helices alone and in combination to alanine and assessed the physical association between these mutants and Vif.

As shown in **Figure SC1-2A**, mutation of the remainder of the α4 helix, excluding E324, failed to render A3F resistant to Vif. In contrast, mutation of the polar residues in the α3 helix to alanine yielded a Vif-resistant A3F variant with a phenotype similar to that of the E324A mutant. Thus, determinants of Vif susceptibility may be found in both the α4 helix and α3 helix of A3F. This is consistent with a contemporaneous report demonstrating that E289, an α3 structural neighbor of E324, is also a single amino acid determinant of A3F susceptibility to Vif (Smith and Pathak, 2010). Although this report indicated that the E289A variant was still somewhat susceptible to Vif, the stronger phenotype observed in the helix mutants described here may indicate a cooperative effect of nearby surface residues such as R293.

We next asked whether physical association between A3F and Vif was interrupted by any of the helix mutants shown in **Figure SC1-2A**. Consistent with a role in Vif

binding, mutation of all polar surface residues in the A3F α4 helix except for E324 resulted in a reproducible decrease in coimmunoprecipitation of A3F by Vif of approximately 50% despite the fact that this mutant was functionally Vif-sensitive (**Figures A1-2B** and **A1-2C**). Curiously, however, when E324A was added to this background, Vif association with A3F was restored to wildtype levels despite the functional resistance of this mutant (**Figure SC1-2A**), and association of the similarly functionally Vif-resistant α3 mutant with Vif was not notably altered. We lack a satisfactory explanation for this phenomenon, but it implies a potential nonequivalence between physical association and functional sensitivity. Another example of this disconnect is the Vif 40-YRHHY-44 region, wherein single amino acid changes at positions 40 and 43 among others yield strong loss of function phenotypes in single-cycle neutralization of A3G as well as loss of coimmunoprecipitation between Vif and A3G (see **Table 1-2**). Despite these strong single-cycle phenotypes, neither the H43N nor the Y40A mutants that we have tested in a proviral context are substantially impaired in their ability to spread in the presence of untagged, stably transfected A3G (See **Supplementary Chapter 2**).

Alternatively, it is possible that some technical aspect of the immunoprecipitation methods described in Chapter 3, such as the C-terminal tags on A3F (V5) and Vif (HA) may somehow alter the immunoprecipitation experiment despite the fact that they faithfully reproduce the functional Vif-A3F interaction as evidenced by more physiologic conditions such as single-cycle titration with untagged A3F and A3F QE323-324EK and spreading infection with the full-length provirus in the presence of cells stably transfected with untagged A3F or the QE323-324EK variant (**Figure 3-2** and **Chapter 5**). We would

argue, however, that positive detection of an interaction is stronger evidence than loss thereof, and more to the point, we speculate that where papers in the field differ on gain or loss of immunoprecipitation [*e.g.* (Xu *et al.*, 2004) vs. (Bogerd *et al.*, 2004; Mangeat *et al.*, 2004; Schröfelbauer *et al.*, 2004) and (Chen *et al.*, 2009; Dang *et al.*, 2009) vs. (Chen *et al.*, 2009; Dang *et al.*, 2009) and **Chapter 2** data not shown] these interpretations are largely a matter of degree. Most authors reporting loss of binding still detect some interaction, and anything from tag identity to buffer composition to the exact mutant employed could account for discrepancies in attempted quantification [*e.g.* (Chen *et al.*, 2009; Dang *et al.*, 2009; Tian *et al.*, 2006)]. The more important point, which we emphasize in our work, is functional interaction or lack thereof.

Functional validation of mutants designated for structural study – As a collaboration with the Matsuo lab at the University of Minnesota and later also with the Schiffer lab at the University of Massachusetts, we have sought to mutagenize the A3F C-terminus to derive a version more amenable to structural study than the wildtype protein. Mutants were initially characterized for enhanced deaminase activity and/or its correlate protein solubility as has been previously reported [*e.g.* (Chen *et al.*, 2008; Chen *et al.*, 2007)]. These efforts led to a complement of deaminase activity enhancing mutations termed A3F 7x, containing mutations Y196D, C259A, F302K, W310K, Y314A, Q315A and F363D. To validate this mutant, we carried out single-cycle titrations of increasing amounts of Vif with a constant amount of A3F 7x and assessed the relative infectivities and A3F stability at each condition. To our surprise, not only was the A3F 7x more

298

restrictive than wildtype A3F, it was also somewhat more Vif-responsive than the wildtype protein (data not shown).

Within this A3F 7x mutant construct, we tested the effect of a number of additional candidate solubility mutations, some of which went on to be incorporated into subsequent generations of mutants made with the intention of rendering the A3F C-terminus more soluble. As shown in **Figure SC1-3A**, these were inconsequential where the Vif interaction is concerned. Additional titration experiments of several intermediate mutants termed 7x, 9x and 11x' also yielded no change in restriction and Vif sensitivity (**Figure SC1-3B**).

The final mutant, A3F 11x, which differs from 11x' only at residue 363, has now been successfully crystallized by the Schiffer lab and will be submitted for publication in the near future. To test this mutant functionally, we again carried out restriction and Vif sensitivity single-cycle infectivity analyses, which demonstrated that A3F 11x is Vif-sensitive and that this sensitivity can be blocked by the Vif-resistance mutation E324K first reported in **Chapter 4** (data not shown). In that experiment, we included Vif G71D in an attempt to further validate the Vif interaction surface of the A3F 11x mutant. While the loss of function against Vif-responsive A3F 11x associated with the G71D mutation was more prominent in the 11x background than in the wildtype protein, the gain of function against A3F 11x E324K was less apparent, particularly by Western blot. It may be, therefore, that the A3F 11x variant is slightly altered in its Vif interaction region, consistent with the presence of changes in the α4 helix (YQ314-315AA). We are currently investigating whether reversion of these specific amino acids to wildtype may

improve the agreement between what we have observed in the wildtype protein in **Chapters 3** and **5** and what we describe here in 11x context.

In summary, despite extensive coverage of surface residues in the A3F C-terminus, critical determinants of Vif susceptibility are tightly localized to two structural neighbors, E324 of the α4 helix and E289 of the α3 helix. It remains possible, however, that less conservative mutations at certain residues and/or the mutation of regions not covered in my work may also affect the Vif-A3F interaction.

A

```
A3F 324 - E Q S L S R L G E Q Y D T D - 311
A3C 141 - E Q S L S R L G E Q Y C P Y - 128
A3DE 337 - E Q S L S C L G E Q Y D T D
                        α4 Helix        Loop
        A3F
        E324K
```

B
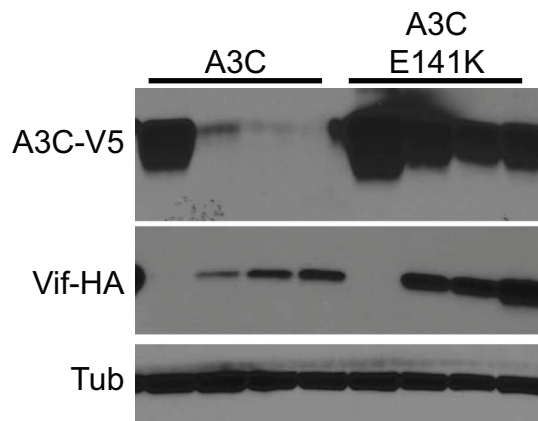
Figure SC1-1: Functional conservation of the Vif interaction site in A3C. **(A)** An alignment showing the conservation in residues of the α4 helix in A3F, A3C and A3D. **(B)** A Vif titration experiment in which cotransfection of a constant amount of A3C or A3C E141K with an increasing amount of Vif results in a decrease in steady state levels of the former but not the latter.
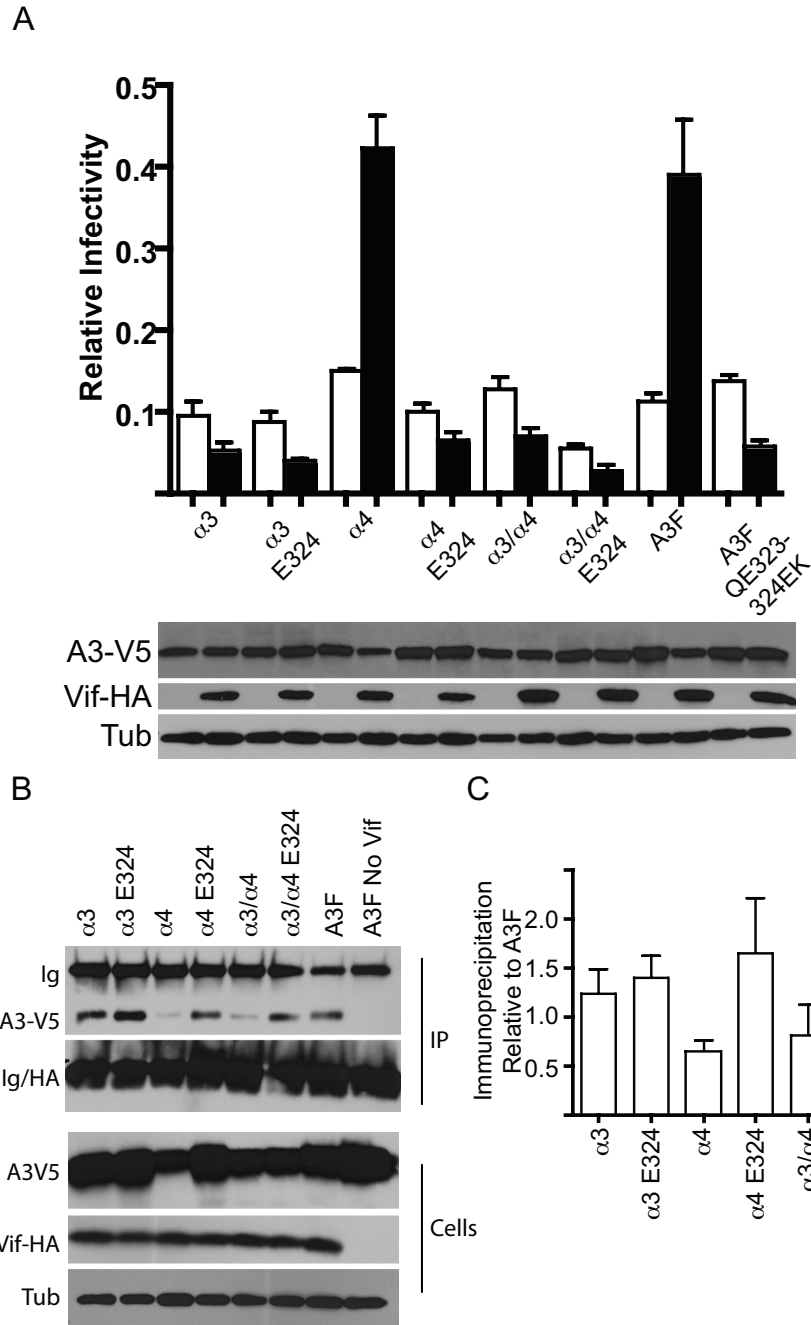
Figure SC1-2: Effect of alanine mutations at α3 and α4 surface residues on Vif susceptibility.

Figure SC1-2: Effect of alanine mutations at α3 and α4 surface residues on Vif susceptibility. (A) A3F containing alanine mutations at all of the polar surface residues in the indicated helix (α4 = S320, E316, D313, D311, +/- E324; α3 = E282, E286, E289, R293, +/- E324) were cotransfected with proviral and Vif-expression plasmids to assess the sensitivity of a given mutant to Vif-mediated degradation and infectivity recovery. **(B)** Immunoprecipitation experiments using the same collection of mutations analyzed in **Figure SC1-2A**. Procedures were as described in **Chapter 3**. **(C)** ImageJ quantification of the relative immunoprecipitation of the indicated A3F variants by Vif. Data derive from four independent experiments.
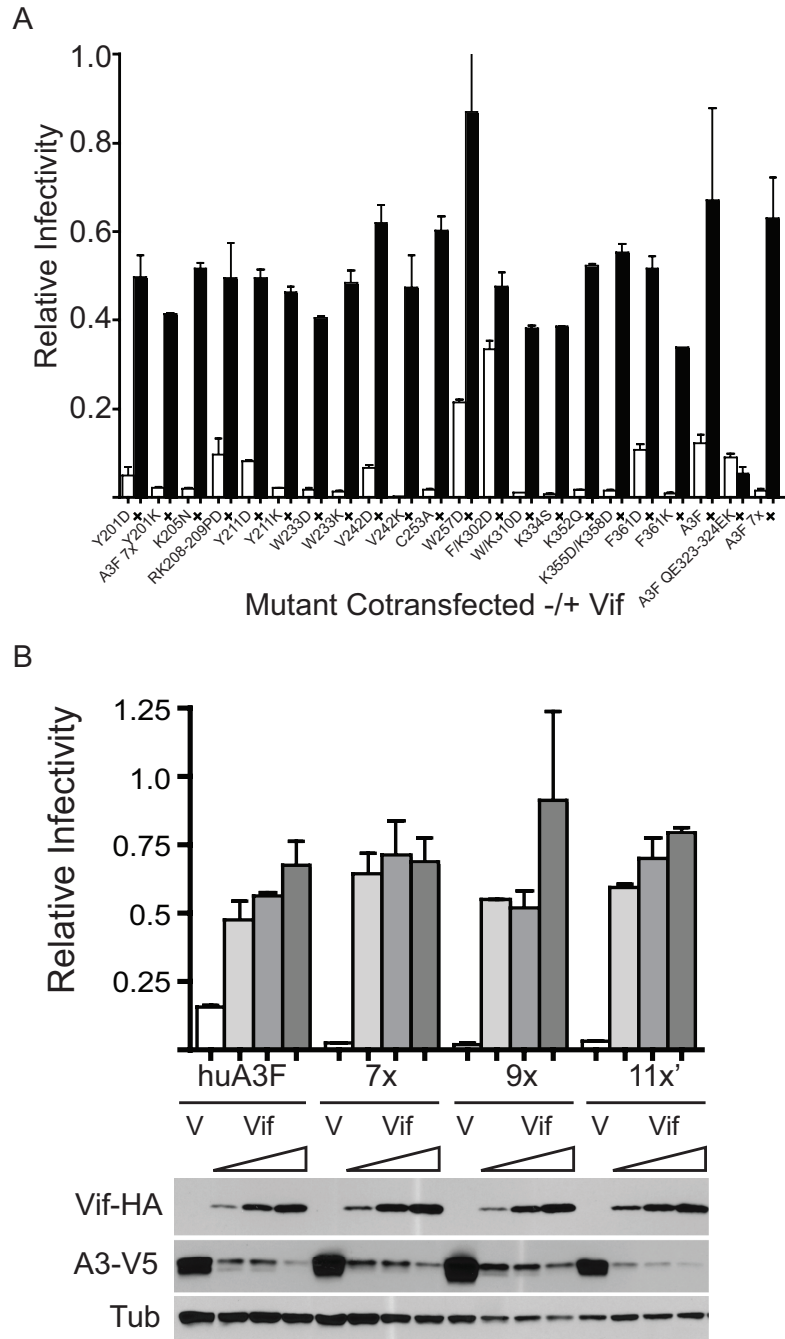
Figure SC1-3: Restriction capacity and Vif sensitivity of A3F mutants derived for structural study.

Figure SC1-3: Restriction capacity and Vif sensitivity of A3F mutants derived for structural study. (A) Single-cycle infectivity of mutations made in A3F 7x context for the derivation of later generations of structural mutants. Methods were similar to those described in **Chapters 2-5**. **(B)** Vif titration experiments demonstrating the HIV restriction capacity and Vif susceptibility of the indicated generational mutants. 7x = Y196D, C259A, F302K, W310K, Y314A, Q315A and F363D; 9x = 7x plus K355D, K358D; 11x = 9x plus H247G, C248R. 11x' in **(B)** has reverted D363 back to the wildtype F363. In both **(A)** and **(B)**, data shown are the mean and standard deviation of duplicate infection events.

Supplementary Chapter 2: Catalytic Activity of APOBEC3F Is Required for HIV

Restriction

Foreword

This supplementary chapter includes data from the manuscript in preparation of the same name concerning the requirement for APOBEC3 catalytic activity in the restriction of HIV by A3F and A3G.

Albin, JS, Brown WL and RS Harris. "Catalytic activity of APOBEC3F is required for HIV restriction." *in preparation*

The reader will recall that this requirement was taken for granted in assessing the encapsidation of A3G in **Chapter 6**, but some authors have proposed that catalytic activity may be dispensable for the ability of A3F to restrict. Here, we weigh in on this question using, for the first time, T cell lines engineered to stably expressed catalytically inactive A3F, thus bypassing the complications associated with transient overexpression.

Summary

APOBEC3 proteins are DNA cytosine deaminases that restrict the replication of diverse retroelements, including human immunodeficiency virus type 1 (HIV) deficient in the counterdefense protein Vif. Among the seven human APOBEC3 proteins, the best studied restrictors of HIV are APOBEC3F (A3F) and APOBEC3G (A3G). The mechanism by which A3F and A3G restrict HIV, however, remains a subject of great interest. While most studies support a role for deaminase activity in the restriction of HIV, a number of reports have proposed accessory mechanisms of restriction that may function in one of several deaminase-independent fashions. Here, we address the capacity of A3F to restrict via deaminase-independent mechanisms by mounting spreading infections in T cells stably transfected with A3F, A3G or deaminase-deficient mutants E251Q and E259Q, respectively. While A3F E251Q induces a slight delay in the replication of Vif-deficient HIV relative both to the analogous A3G mutant E259Q and to APOBEC3-deficient cell lines, it is unable to cause long-term suppression of virus replication, unlike deaminase-proficient wildtype A3F and A3G. We therefore conclude that the primary mechanism of A3F-mediated restriction of HIV that forces the virus to maintain Vif function is deaminase-dependent.

APOBEC3 proteins restrict HIV in the absence of the viral accessory protein Vif. Most authors agree that the deaminase activity of APOBEC3 proteins is an important part of the restriction mechanism, a model wherein the deamination of single-stranded minus-strand cDNA during reverse transcription results in deleterious hypermutation of the viral genome and ultimately the inhibition of viral replication (Harris *et al.*, 2003a; Mangeat *et al.*, 2003; Yu *et al.*, 2004b; Zhang *et al.*, 2003). Despite this consensus, many authors have proposed diverse deaminase-independent mechanisms that may also operate in the restriction of HIV including direction physical inhibition of reverse transcriptase processivity by RNA-bound APOBEC3, inhibition of tRNA priming and the inhibition of strand transfer events (Bishop *et al.*, 2006; Bishop *et al.*, 2008; Guo *et al.*, 2006; Holmes *et al.*, 2007; Iwatani *et al.*, 2007; Li *et al.*, 2007; Luo *et al.*, 2007; Mbisa *et al.*, 2007; Mbisa *et al.*, 2010; Newman *et al.*, 2005; Shindo *et al.*, 2003; Yang *et al.*, 2007b).

More recently, several groups have reported that these deaminase-independent mechanisms may represent overexpression artifacts since, when one stably expresses an A3G catalytic mutant or titrates back the levels of transiently expressed protein, the restrictive effects of deaminase-deficient A3G prove incapable of achieving more than modest if any restriction of HIV (Browne *et al.*, 2009; Miyagi *et al.*, 2007; Schumacher *et al.*, 2008). This same story has played out in less mechanistic detail with the related restriction factor A3F, which was initially proposed to have more prominent deaminase-independent effects than A3G (Holmes *et al.*, 2007). When utilized in a stable expression system, however, an A3F variant with mutations in the zinc-coordinating cysteines, C280S/C283S, failed to restrict HIV as had A3G catalytic mutants before it (Miyagi *et al.*, 2010). Curiously, however, this report also failed to detect restriction by wildtype

A3F stably expressed in HeLa cells, which contrasts with the restrictive activities we have found attributable to A3F in no fewer than 17 independently-derived T cell lines stably transfected with A3F (Albin *et al.*, 2010a; Albin *et al.*, 2010b; Haché *et al.*, 2008; Hultquist *et al.*, 2011).

To determine whether deaminase-deficient A3F is capable of restricting HIV in continuously cultured T cell lines, we stably transfected a permissive cell line, SupT11, with wildtype A3F, A3G or catalytic mutant derivatives thereof. We specifically elected to use mutants of the conserved catalytic glutamate in A3F (E251Q) or A3G (E259Q) since these show little or no discernible encapsidation defect relative to the wildtype protein [*e.g.* (Holmes *et al.*, 2007; Schumacher *et al.*, 2008) and **Figure SC2-1**]. Expression levels of A3F and A3G in all lines used were similar between wildtype and catalytic mutant lines (**Figure SC2-2**). We then initiated infections of all of these lines in parallel at MOIs of 0.01 and 0.05 with either wildtype HIV$_{IIIB}$ or a Vif-deficient derivative containing tandem stops at codons 26-27 of *vif*. Representative spreading infection curves are shown in **Figure SC2-3**. As shown, while stable expression of either A3F or A3G at levels comparable to those occurring in naturally nonpermissive cells restricts the spread of HIV for the duration of the experiment, Vif-deficient viruses peak within approximately two weeks on cell lines expressing catalytic mutant derivatives of A3F or A3G with a slightly longer delay induced by the former. Although we did not carry out these experiments any longer than is shown in **Figure SC2-3**, we have previously noted by using similar methods of infection with lines stably expressing A3F that this restriction can extend for long periods of time, preventing the spread of Vif-

deficient HIV for at least three months in the absence of the acquisition of genetic resistance such as repair of the *vif* open reading frame (Albin *et al.*, 2010a).

Despite the inability of catalytically inactive A3F and A3G to restrict on a par with the wildtype proteins, we did note a small but consistent delay in the spread of Vif-deficient HIV relative to cells stably transfected with a vector control (data not shown). It is therefore formally possible that A3F may display some minor but appreciable deaminase-independent effects relative to A3G, as initially proposed (Holmes *et al.*, 2007). It is important to emphasize, however, is that this putative activity is incapable of inducing more than a modest delay and cannot account for the long-term duration of restriction that we have previously observed with A3F in the absence of genetic resistance (Albin *et al.*, 2010a).

Others have proposed that the levels of A3F in primary cells may be incapable of restricting HIV (Mulder *et al.*, 2010). In this study, the authors utilized a Vif mutant, W11R, that is capable of neutralizing A3G but not A3F to ask whether expression levels of the latter in *ex vivo* peripheral blood mononuclear cells from several donors are incapable of restricting viral spread. We were curious to test this mutant, however, because unrelated studies in our lab aimed at characterizing *vif* alleles that arise during long-term viral passages in the presence of APOBEC3 proteins had cast some doubt on the concept of selectively susceptible Vif mutants (see **Chapter 5**). Specifically, we had noted that mutants of the YRHHY region often implicated in the interaction of HIV Vif with A3G were largely replication-competent in the presence of stably expressed A3G despite their strong selectively susceptible single-cycle phenotypes [(Russell and Pathak, 2007) and Albin & Harris, unpublished data]. Similarly, others had noted that long-term

passage of YRHHY mutants in naturally nonpermissive CEM cells fails to select for reversion of the mutated Vif residues, suggesting lack of selective pressure and, by extension, intact anti-A3G function in that spreading infection system (Russell *et al.*, 2009a).

To test the selective susceptibility of the W11R mutant as well as select passage-derived mutants of the putative A3F- and A3G-interacting regions of Vif ([14]DRMR[17] and [40]YRHHY[44], respectively), we carried out spreading infections in the presence of A3F, A3G and naturally nonpermissive CEM2n cells using wildtype and Vif-deficient HIV_{IIIB} as before as well as mutant W11R. In these experiments, we also included three mutants of the relevant regions that we had identified incidentally in unrelated long-term passage experiments: R15G, H43N and H43N/E117K. Mutation of each of these residues has been previously shown to ablate Vif anti-A3F (residues 11 and 15) or anti-A3G (residue 43) activity in transient overexpression systems (see **Table 1-2**). Consistent with those overexpression data, each of the mutants predicted to be selectively susceptible to A3F was restricted by stable expression of A3F in SupT11 cells as well as by the levels of A3F expressed in naturally nonpermissive CEM2n cells but not by A3G (**Figure SC2-4**). In contrast, mutant H43N or H43N in combination with an additional passenger mutation E117K was capable of spreading efficiently with perhaps a modest delay in the presence of both A3G stably expressed in SupT11 cells as well as the levels of A3G expressed in CEM2n cells (**Figure SC2-4**). These results mirrored those we had previously observed with two independent mutants of the [14]DRMR[17] (D14A) and [40]YRHHY[44] (Y40A) regions at a time when we, too, were interested in probing the contributions of different APOBEC3 proteins in naturally nonpermissive cells using selectively susceptible Vif

variants (Albin and Harris, unpublished data). Together, our data suggest that, while mutants of the ^{14}DRMR^{17} region are indeed selectively susceptible to A3F, the phenotypes associated with mutants of the ^{40}YRHHY^{44} region are at least partially dependent on the experimental system used, showing strong effects in transient overexpression systems and poor effects in stable expression systems. This contrasts with mutants of Vif residue 26, which are consistently selectively defective for the neutralization of A3G in both transient and stable expression systems (Albin *et al.*, 2010a; Chen *et al.*, 2009; Dang *et al.*, 2009).

We therefore question whether the loss-of-function mutants previously identified as critical A3F- and A3G-binding regions are, in fact, A3F and A3G interacting regions. The inability of ^{40}YRHHY^{44} mutants to faithfully reproduce their susceptibility phenotypes outside of transient overexpression systems suggests that their effects may be more complicated than a simple loss of anti-A3G activity. Moreover, while ^{14}DRMR^{17} mutants are phenotypically consistent in their loss of anti-A3F function, it is curious that mutation of this would-be A3F-selective region can result in the neutralization of A3G mutants containing alterations in the putative Vif binding residue D128K (Schröfelbauer *et al.*, 2006). It is possible, therefore, that the losses of function associated with each region may represent something other than the simple loss of direct binding proposed.

Despite the unexpected growth of ^{40}YRHHY^{44} mutants in the presence of restrictive levels of A3G, we confirmed in our system that the W11R mutant displays selective susceptibility to A3F. It may therefore be accurate that the levels of A3F in *ex vivo* samples from the donors previously described are unable to restrict HIV, although data concerning the ability of similar mutants selectively susceptible to A3G rather than

A3F were not reported (Mulder *et al.*, 2010). Nevertheless, we note that both the W11R and the R15G mutants utilized in the present studies fail to spread in naturally nonpermissive cells. If one accepts these mutants as truly selectively susceptible to A3F, then, that would suggest that the levels of A3F present in naturally nonpermissive cells are sufficient to induce indefinite restriction of the kind we've previously reported via stable expression of A3F. In that case, it would seem appropriate to ask not whether A3F restricts in primary cells, but rather whether the expression of any APOBEC3 protein is adequate to prevent viral spread in primary cells, particularly *in vivo*.

In conclusion, we have shown here that the restriction of HIV by both A3F and A3G requires intact deaminase activity. This suggests that the predominant driving force for the maintenance of Vif function by the virus centers not on the comparatively minor effects potentially imposed by deaminase-independent mechanisms of restriction, but rather on the deaminase-dependent components previously described. We further note that not all Vif mutants thought to be selectively susceptible to one APOBEC3 protein or another yield similar phenotypes in transient overexpression and spreading infection systems. It is therefore possible that the ever-expanding literature on loss of function Vif mutants may be of limited utility in defining the regions of Vif that directly interact with A3F and A3G since, by nature, loss of function can indicate many things other than loss of direct binding [reviewed in (Albin, 2010; Smith *et al.*, 2009)].
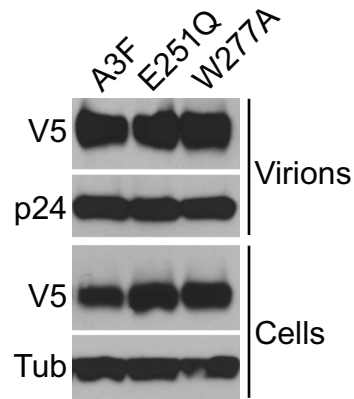
Figure SC2-1: Encapsidation of A3F catalytic mutants. 293T cells were transfected with 200 ng of V5-tagged wildtype A3F, A3F E251Q or another deaminase-deficient A3F derivative, W277A, along with 1.6 μg Vif-deficient HIV$_{IIIB}$. Producer cells and supernatants were harvested two days post-transfection, and supernatants were spun through 20% sucrose cushions and lysed. Western blots for particle-associated V5 and p24 as well as cell-associated V5 and tubulin are shown.
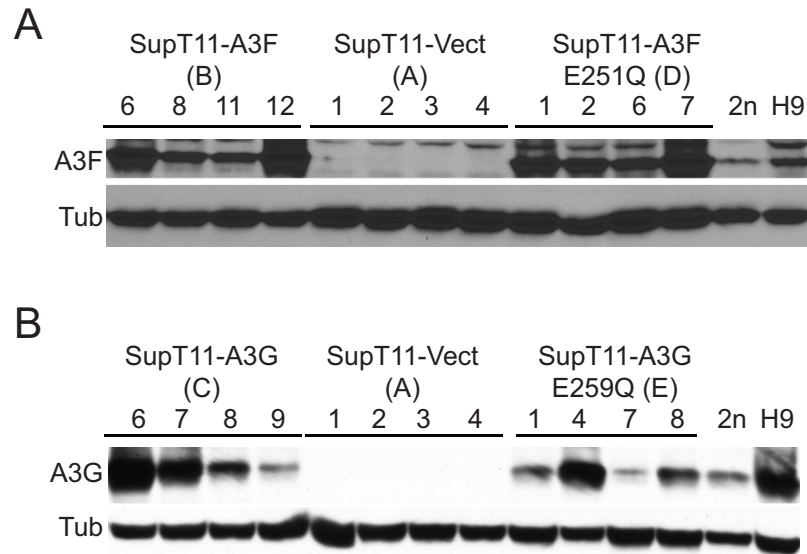
Figure SC2-2: APOBEC3 expression levels in the cell lines used in this study. (A) A3F levels found in SupT11 cells stably transfected with untagged A3F, A3F E251Q or a vector control in comparison with naturally nonpermissive lines CEM2n and H9. **(B)** A3G levels found in SupT11 cells stably transfected with untagged A3G, A3G E259Q or a vector control in comparison with naturally nonpermissive lines CEM2n and H9.
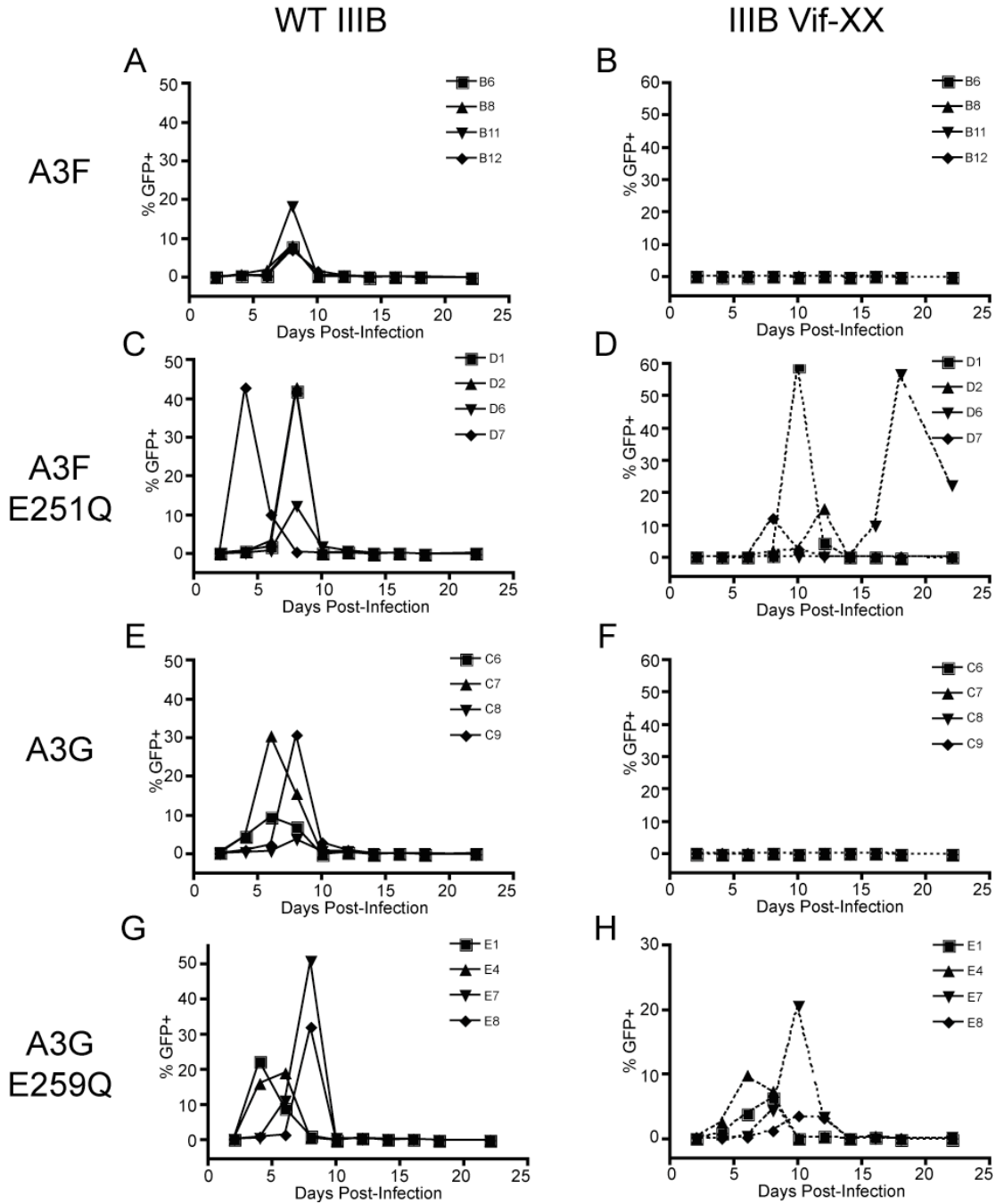
Figure SC2-3: Deaminase activity is required for restriction of HIV by both A3F and A3G.

Figure SC2-3: Deaminase activity is required for restriction of HIV by both A3F and A3G. Representative spreading infection curves infected at MOI 0.05 from one of three, independent experiments demonstrating the inability of catalytic mutants of A3F and A3G to efficiently restrict HIV. **(A-B)** Wildtype and Vif-XX variants of HIV_{IIIB} on four independent cell lines expressing wildtype A3F. **(C-D)** Wildtype and Vif-XX variants of HIV_{IIIB} on four independent cell lines expressing A3F E251Q. **(E-F)** Wildtype and Vif-XX variants of HIV_{IIIB} on four independent cell lines expressing wildtype A3G. **(G-H)** Wildtype and Vif-XX variants of HIV_{IIIB} on four independent cell lines expressing A3G E259Q. Note that each individual cell line is a stably transfected clone established by limiting dilution. Different cell lines expressing the same APOBEC3 protein are therefore susceptible to clonal variation but are graphed together to demonstrate that all cell lines qualitatively recapitulate growth versus no growth of a given virus. All viral cultures that peaked also subsequently killed all host cells in a given culture, whereas those that did not peak maintained healthy cell growth throughout.
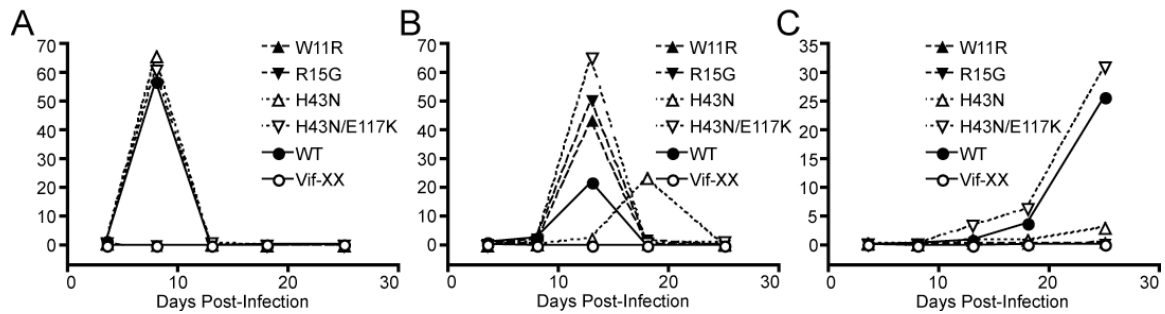
Figure SC2-4: Mutation of the putative A3F-interacting region of Vif but not the putative A3G-interacting region of Vif yields viruses selectively susceptible to the predicted APOBEC3 protein. (A) Growth of wildtype, Vif-XX and the indicated Vif mutant derivatives of HIV$_{IIIB}$ in the presence of A3F. Results are representative of 2-3 independent experiments on four independent cell lines. **(B)** Growth of wildtype, Vif-XX and the indicated Vif mutants in the presence of A3G. Results are representative of 2-3 independent experiments on four independent cell lines. **(C)** Growth of wildtype, Vif-XX and the indicated Vif mutants in naturally nonpermissive CEM2n cells. Results are representative of three independent experiments.

References

UNAIDS, J.U.N.P.o.H.A. (2010a). UNAIDS global report fact sheet 2010.

UNAIDS, J.U.N.P.o.H.A. (2010b). UNAIDS report on the global AIDS epidemic 2010.

UNAIDS, J.U.N.P.o.H.A. (2012). Global AIDS epidemic facts and figures.

Abudu, A., Takaori-Kondo, A., Izumi, T., Shirakawa, K., Kobayashi, M., Sasada, A., Fukunaga, K., and Uchiyama, T. (2006). Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. Curr Biol *16*, 1565-1570.

Aguiar, R.S., Lovsin, N., Tanuri, A., and Peterlin, B.M. (2008). Vpr.A3A chimera inhibits HIV replication. The Journal of Biological Chemistry *283*, 2518-2525.

Albin, J.S., and R.S. Harris (2010). Interactions of host APOBEC3 restriction factors with HIV *in vivo*: implications for therapeutics. Expert Reviews in Molecular Medicine *12*, 1-26.

Albin, J.S., Haché, G., Hultquist, J.F., Brown, W.L., and Harris, R.S. (2010a). Long-term Restriction by APOBEC3F Selects Human Immunodeficiency Virus Type 1 Variants with Restored Vif Function. Journal of Virology *84*, 10209-10219.

Albin, J.S., LaRue, R.S., Weaver, J.A., Brown, W.L., Shindo, K., Harjes, E., Matsuo, H., and Harris, R.S. (2010b). A single amino acid in human APOBEC3F alters susceptibility to HIV Vif. The Journal of Biological Chemistry *285*, 40785-40792.

Alce, T.M., and Popik, W. (2004). APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV Gag nucleocapsid protein. The Journal of Biological Chemistry *279*, 34083-34086.

Amoedo, N.D., Afonso, A.O., Cunha, S.M., Oliveira, R.H., Machado, E.S., and Soares, M.A. (2011). Expression of APOBEC3G/3F and G-to-A Hypermutation Levels in HIV-Infected Children with Different Profiles of Disease Progression. PLoS One *6*, e24118.

An, P., Bleiber, G., Duggal, P., Nelson, G., May, M., Mangeat, B., Alobwede, I., Trono, D., Vlahov, D., Donfield, S.*, et al.* (2004). APOBEC3G genetic variants and their influence on the progression to AIDS. Journal of Virology *78*, 11070-11076.

An, P., Johnson, R., Phair, J., Kirk, G.D., Yu, X.F., Donfield, S., Buchbinder, S., Goedert, J.J., and Winkler, C.A. (2009). APOBEC3B deletion and risk of HIV acquisition. J Infect Dis *200*, 1054-1058.

Arnesen, T. (2011). Towards a functional understanding of protein N-terminal acetylation. PLoS Biol *9*, e1001074.

Auclair, J.R., Green, K.M., Shandilya, S., Evans, J.E., Somasundaran, M., and Schiffer, C.A. (2007). Mass spectrometry analysis of HIV Vif reveals an increase in ordered structure upon oligomerization in regions necessary for viral infectivity. Proteins *69*, 270-284.

Autore, F., Bergeron, J.R., Malim, M.H., Fraternali, F., and Huthoff, H. (2010). Rationalisation of the differences between APOBEC3G structures from crystallography and NMR studies by molecular dynamics simulations. PLoS One *5*, e11515.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). *In vivo* half-life of a protein is a function of its amino-terminal residue. Science *234*, 179-186.

Batonnet, S., Leibovitch, M.P., Tintignac, L., and Leibovitch, S.A. (2004). Critical role for lysine 133 in the nuclear ubiquitin-mediated degradation of MyoD. The Journal of Biological Chemistry *279*, 5413-5420.

Beale, R.C., Petersen-Mahrt, S.K., Watt, I.N., Harris, R.S., Rada, C., and Neuberger, M.S. (2004). Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra *in vivo*. J Mol Biol *337*, 585-596.

Berger, A., Munk, C., Schweizer, M., Cichutek, K., Schule, S., and Flory, E. (2010). Interaction of Vpx and apolipoprotein B mRNA-editing catalytic polypeptide 3 family member A (APOBEC3A) correlates with efficient lentivirus infection of monocytes. The Journal of Biological Chemistry *285*, 12248-12254.

Berger, A., Sommer, A.F., Zwarg, J., Hamdorf, M., Welzel, K., Esly, N., Panitz, S., Reuter, A., Ramos, I., Jatiani, A.*, et al.* (2011a). SAMHD1-Deficient CD14+ Cells from Individuals with Aicardi-Goutières Syndrome Are Highly Susceptible to HIV Infection. PLoS Pathog *7*, e1002425.

Berger, G., Durand, S., Fargier, G., Nguyen, X.N., Cordeil, S., Bouaziz, S., Muriaux, D., Darlix, J.L., and Cimarelli, A. (2011b). APOBEC3A is a specific inhibitor of the early phases of HIV infection in myeloid cells. PLoS Pathog *7*, e1002221.

Bergeron, J.R., Huthoff, H., Veselkov, D.A., Beavil, R.L., Simpson, P.J., Matthews, S.J., Malim, M.H., and Sanderson, M.R. (2010). The SOCS-box of HIV Vif interacts with ElonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex. PLoS Pathog *6*, e1000925.

Berkhout, B., and de Ronde, A. (2004). APOBEC3G versus reverse transcriptase in the generation of HIV drug-resistance mutations. AIDS *18*, 1861-1863.

Bewick, S., Wu, J., Lenaghan, S.C., Yang, R., Zhang, M., and Hamel, W. (2011). The R5 to X4 Coreceptor Switch: A Dead-End Path, or a Strategic Maneuver? : Lessons from a Game Theoretic Analysis. Bull Math Biol *73*, 2339-2356.

Biasin, M., Piacentini, L., Lo Caputo, S., Kanari, Y., Magri, G., Trabattoni, D., Naddeo, V., Lopalco, L., Clivio, A., Cesana, E.*, et al.* (2007). 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 *195*, 960-964.

Binka, M., Ooms, M., Steward, M., and Simon, V. (2012). The activity spectrum of Vif from multiple HIV subtypes against APOBEC3G, APOBEC3F, and APOBEC3H. Journal of Virology *86*, 49-59.

Bishop, K.N., Holmes, R.K., and Malim, M.H. (2006). Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. Journal of Virology *80*, 8450-8458.

Bishop, K.N., Holmes, R.K., Sheehy, A.M., Davidson, N.O., Cho, S.J., and Malim, M.H. (2004). Cytidine deamination of retroviral DNA by diverse APOBEC proteins. Curr Biol *14*, 1392-1396.

Bishop, K.N., Verma, M., Kim, E.Y., Wolinsky, S.M., and Malim, M.H. (2008). APOBEC3G inhibits elongation of HIV reverse transcripts. PLoS Pathog *4*, e1000231.

Bogerd, H.P., and Cullen, B.R. (2008). Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation. RNA *14*, 1228-1236.

320

Bogerd, H.P., Doehle, B.P., Wiegand, H.L., and Cullen, B.R. (2004). A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor. Proceedings of the National Academy of Sciences of the United States of America *101*, 3770-3774.

Bogerd, H.P., Tallmadge, R.L., Oaks, J.L., Carpenter, S., and Cullen, B.R. (2008). Equine infectious anemia virus resists the antiretroviral activity of equine APOBEC3 proteins through a packaging-independent mechanism. Journal of Virology *82*, 11889-11901.

Bogerd, H.P., Wiegand, H.L., Doehle, B.P., and Cullen, B.R. (2007). The intrinsic antiretroviral factor APOBEC3B contains two enzymatically active cytidine deaminase domains. Virology *364*, 486-493.

Bogerd, H.P., Wiegand, H.L., Doehle, B.P., Lueders, K.K., and Cullen, B.R. (2006a). APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. Nucleic Acids Research *34*, 89-95.

Bogerd, H.P., Wiegand, H.L., Hulme, A.E., Garcia-Perez, J.L., O'Shea, K.S., Moran, J.V., and Cullen, B.R. (2006b). Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proceedings of the National Academy of Sciences of the United States of America *103*, 8780-8785.

Bohn, M.F., Shandilya, S.M., Albin, J.S., Kono, T., Carpenter, M.A., Anderson, B.D., Davis, A.N., Zhang, J.Y., Lu, Y.J., Somasundaran, M.*, et al.* Crystal Structure of the HIV Vif binding and catalytically active domain of APOBEC3F. *In preparation*.

Borman, A.M., Quillent, C., Charneau, P., Dauguet, C., and Clavel, F. (1995). Human immunodeficiency virus type 1 Vif- mutant particles from restrictive cells: role of Vif in correct particle assembly and infectivity. Journal of Virology *69*, 2058-2067.

Bourara, K., Liegler, T.J., and Grant, R.M. (2007). Target cell APOBEC3C can induce limited G-to-A mutation in HIV. PLoS Pathog *3*, 1477-1485.

Breaker, R.R. (2004). Natural and engineered nucleic acids as tools to explore biology. Nature *432*, 838-845.

Britan-Rosich, E., Nowarski, R., and Kotler, M. (2011). Multifaceted counter-APOBEC3G mechanisms employed by HIV Vif. J Mol Biol *410*, 1065-1076.

Browne, E.P., Allers, C., and Landau, N.R. (2009). Restriction of HIV by APOBEC3G is cytidine deaminase-dependent. Virology *387*, 313-321.

Cagliani, R., Riva, S., Fumagalli, M., Biasin, M., Caputo, S.L., Mazzotta, F., Piacentini, L., Pozzoli, U., Bresolin, N., Clerici, M.*, et al.* (2011). A positively selected APOBEC3H haplotype is associated with natural resistance to HIV infection. Evolution *65*, 3311-3322.

Cai, H., Bloom, L.B., Eritja, R., and Goodman, M.F. (1993). Kinetics of deoxyribonucleotide insertion and extension at abasic template lesions in different sequence contexts using HIV reverse transcriptase. The Journal of Biological Chemistry *268*, 23567-23572.

Cancio, R., Spadari, S., and Maga, G. (2004). Vif is an auxiliary factor of the HIV reverse transcriptase and facilitates abasic site bypass. Biochem J *383*, 475-482.

Carpenter, M.A., Rajagurubandara, E., Wijesinghe, P., and Bhagwat, A.S. (2010). Determinants of sequence-specificity within human AID and APOBEC3G. DNA Repair (Amst) *9*, 579-587.

Carr, J.M., Coolen, C., Davis, A.J., Burrell, C.J., and Li, P. (2008). Human immunodeficiency virus 1 (HIV) virion infectivity factor (Vif) is part of reverse transcription complexes and acts as an accessory factor for reverse transcription. Virology *372*, 147-156.

Carr, J.M., Davis, A.J., Coolen, C., Cheney, K., Burrell, C.J., and Li, P. (2006). Vif-deficient HIV reverse transcription complexes (RTCs) are subject to structural changes and mutation of RTC-associated reverse transcription products. Virology *351*, 80-91.

Cen, S., Guo, F., Niu, M., Saadatmand, J., Deflassieux, J., and Kleiman, L. (2004). The interaction between HIV Gag and APOBEC3G. The Journal of Biological Chemistry *279*, 33177-33184.

Chelico, L., Pham, P., Calabrese, P., and Goodman, M.F. (2006). APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA. Nat Struct Mol Biol *13*, 392-399.

Chelico, L., Prochnow, C., Erie, D.A., Chen, X.S., and Goodman, M.F. (2010). Structural model for deoxycytidine deamination mechanisms of the HIV inactivation enzyme APOBEC3G. The Journal of Biological Chemistry *285*, 16195-16205.

Chelico, L., Sacho, E.J., Erie, D.A., and Goodman, M.F. (2008). A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. The Journal of Biological Chemistry *283*, 13780-13791.

Chen, G., He, Z., Wang, T., Xu, R., and Yu, X.F. (2009). A patch of positively charged amino acids surrounding the human immunodeficiency virus type 1 Vif SLVx4Yx9Y motif influences its interaction with APOBEC3G. Journal of Virology *83*, 8674-8682.

Chen, H., Lilley, C.E., Yu, Q., Lee, D.V., Chou, J., Narvaiza, I., Landau, N.R., and Weitzman, M.D. (2006). APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. Curr Biol *16*, 480-485.

Chen, K.M., Harjes, E., Gross, P.J., Fahmy, A., Lu, Y., Shindo, K., Harris, R.S., and Matsuo, H. (2008). Structure of the DNA deaminase domain of the HIV restriction factor APOBEC3G. Nature *452*, 116-119.

Chen, K.M., Martemyanova, N., Lu, Y., Shindo, K., Matsuo, H., and Harris, R.S. (2007). Extensive mutagenesis experiments corroborate a structural model for the DNA deaminase domain of APOBEC3G. FEBS Lett *581*, 4761-4766.

Chen, R., Le Rouzic, E., Kearney, J.A., Mansky, L.M., and Benichou, S. (2004). Vpr-mediated incorporation of UNG2 into HIV particles is required to modulate the virus mutation rate and for replication in macrophages. The Journal of Biological Chemistry *279*, 28419-28425.

Chiu, Y.L., and Greene, W.C. (2008). The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. Annu Rev Immunol *26*, 317-353.

Cho, S.J., Drechsler, H., Burke, R.C., Arens, M.Q., Powderly, W., and Davidson, N.O. (2006). APOBEC3F and APOBEC3G mRNA levels do not correlate with human

immunodeficiency virus type 1 plasma viremia or CD4+ T-cell count. Journal of Virology *80*, 2069-2072.

Chowers, M.Y., Spina, C.A., Kwoh, T.J., Fitch, N.J., Richman, D.D., and Guatelli, J.C. (1994). Optimal infectivity *in vitro* of human immunodeficiency virus type 1 requires an intact nef gene. Journal of Virology *68*, 2906-2914.

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

Conticello, S.G., Harris, R.S., and Neuberger, M.S. (2003). The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. Curr Biol *13*, 2009-2013.

Conticello, S.G., Thomas, C.J., Petersen-Mahrt, S., and Neuberger, M.S. (2005). Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Mol Biol Evol *22*, 367-377.

Courcoul, M., Patience, C., Rey, F., Blanc, D., Harmache, A., Sire, J., Vigne, R., and Spire, B. (1995). Peripheral blood mononuclear cells produce normal amounts of defective Vif- human immunodeficiency virus type 1 particles which are restricted for the preretrotranscription steps. Journal of Virology *69*, 2068-2074.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Cummins, N.W., Rizza, S.A., and Badley, A.D. (2010). How much gp120 is there? J Infect Dis *201*, 1273-1274; author reply 1274-1275.

Dang, Y., Davis, R.W., York, I.A., and Zheng, Y.H. (2010a). Identification of 81LGxGxxIxW89 and 171EDRW174 domains from human immunodeficiency virus type 1 Vif that regulate APOBEC3G and APOBEC3F neutralizing activity. Journal of Virology *84*, 5741-5750.

Dang, Y., Siew, L.M., Wang, X., Han, Y., Lampen, R., and Zheng, Y.H. (2008a). Human cytidine deaminase APOBEC3H restricts HIV replication. The Journal of Biological Chemistry *283*, 11606-11614.

Dang, Y., Siew, L.M., and Zheng, Y.H. (2008b). APOBEC3G is degraded by the proteasomal pathway in a Vif-dependent manner without being polyubiquitylated. The Journal of Biological Chemistry *283*, 13124-13131.

Dang, Y., Wang, X., Esselman, W.J., and Zheng, Y.H. (2006). Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. Journal of Virology *80*, 10522-10533.

Dang, Y., Wang, X., York, I.A., and Zheng, Y.H. (2010b). Identification of a critical T(Q/D/E)x5ADx2(I/L) motif from primate lentivirus Vif proteins that regulate APOBEC3G and APOBEC3F neutralizing activity. Journal of Virology *84*, 8561-8570.

Dang, Y., Wang, X., Zhou, T., York, I.A., and Zheng, Y.H. (2009). 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. Journal of Virology *83*, 8544-8552.

Das-Bradoo, S., Nguyen, H.D., Wood, J.L., Ricke, R.M., Haworth, J.C., and Bielinsky, A.K. (2010). Defects in DNA ligase I trigger PCNA ubiquitylation at Lys 107. Nat Cell Biol *12*, 74-79; sup pp 71-20.

De Maio, F.A., Rocco, C.A., Aulicino, P.C., Bologna, R., Mangano, A., and Sen, L. (2011). Effect of HIV Vif variability on progression to pediatric AIDS and its association with APOBEC3G and CUL5 polymorphisms. Infect Genet Evol *11*, 1256-1262.

Derse, D., Hill, S.A., Princler, G., Lloyd, P., and Heidecker, G. (2007). Resistance of human T cell leukemia virus type 1 to APOBEC3G restriction is mediated by elements in nucleocapsid. Proceedings of the National Academy of Sciences of the United States of America *104*, 2915-2920.

Deshaies, R.J., and Joazeiro, C.A. (2009). RING domain E3 ubiquitin ligases. Annu Rev Biochem *78*, 399-434.

Desrosiers, R.C., Lifson, J.D., Gibbs, J.S., Czajak, S.C., Howe, A.Y., Arthur, L.O., and Johnson, R.P. (1998). Identification of highly attenuated mutants of simian immunodeficiency virus. Journal of Virology *72*, 1431-1437.

Dewal, M.B., and Firestine, S.M. (2011). Non-peptidic alpha-helical mimetics as protein-protein interaction inhibitors. Curr Med Chem *18*, 2420-2428.

Do, H., Vasilescu, A., Diop, G., Hirtzig, T., Heath, S.C., Coulonges, C., Rappaport, J., Therwath, A., Lathrop, M., Matsuda, F.*, et al.* (2005). Exhaustive genotyping of the CEM15 (APOBEC3G) gene and absence of association with AIDS progression in a French cohort. J Infect Dis *191*, 159-163.

Doehle, B.P., Bogerd, H.P., Wiegand, H.L., Jouvenet, N., Bieniasz, P.D., Hunter, E., and Cullen, B.R. (2006). The betaretrovirus Mason-Pfizer monkey virus selectively excludes simian APOBEC3G from virion particles. Journal of Virology *80*, 12102-12108.

Doehle, B.P., Schafer, A., and Cullen, B.R. (2005a). Human APOBEC3B is a potent inhibitor of HIV infectivity and is resistant to HIV Vif. Virology *339*, 281-288.

Doehle, B.P., Schafer, A., Wiegand, H.L., Bogerd, H.P., and Cullen, B.R. (2005b). Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. Journal of Virology *79*, 8201-8207.

Donahue, J.P., Vetter, M.L., Mukhtar, N.A., and D'Aquila, R.T. (2008). The HIV Vif PPLP motif is necessary for human APOBEC3G binding and degradation. Virology *377*, 49-53.

Douglas, J.L., Gustin, J.K., Viswanathan, K., Mansouri, M., Moses, A.V., and Fruh, K. (2010). The great escape: viral strategies to counter BST-2/tetherin. PLoS Pathog *6*, e1000913.

Duggal, N.K., Malik, H.S., and Emerman, M. (2011). Positive selection of Apobec3DE in chimpanzees has driven breadth in anti-viral activity. Journal of Virology *85*, 11361-11371.

Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A.J., Heidmann, T., and Schwartz, O. (2005). APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature *433*, 430-433.

Esnault, C., Millet, J., Schwartz, O., and Heidmann, T. (2006). Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. Nucleic Acids Research *34*, 1522-1531.

Feng, Y., and Chelico, L. (2011). Intensity of deoxycytidine deamination of HIV proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain. The Journal of Biological Chemistry *286*, 11415-11426.

Fisher, A.G., Ensoli, B., Ivanoff, L., Chamberlain, M., Petteway, S., Ratner, L., Gallo, R.C., and Wong-Staal, F. (1987). The sor gene of HIV is required for efficient virus transmission *in vitro*. Science *237*, 888-893.

Fourati, S., Malet, I., Binka, M., Boukobza, S., Wirden, M., Sayon, S., Simon, A., Katlama, C., Simon, V., Calvez, V.*, et al.* (2010). Partially active HIV Vif alleles facilitate viral escape from specific antiretrovirals. AIDS *24*, 2313-2321.

Friew, Y.N., Boyko, V., Hu, W.S., and Pathak, V.K. (2009). Intracellular interactions between APOBEC3G, RNA, and HIV Gag: APOBEC3G multimerization is dependent on its association with RNA. Retrovirology *6*, 56.

Furukawa, A., Nagata, T., Matsugami, A., Habu, Y., Sugiyama, R., Hayashi, F., Kobayashi, N., Yokoyama, S., Takaku, H., and Katahira, M. (2009). Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G. Embo J *28*, 440-451.

Gabuzda, D.H., Lawrence, K., Langhoff, E., Terwilliger, E., Dorfman, T., Haseltine, W.A., and Sodroski, J. (1992). Role of Vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. Journal of Virology *66*, 6489-6495.

Gabuzda, D.H., Li, H., Lawrence, K., Vasir, B.S., Crawford, K., and Langhoff, E. (1994). Essential role of vif in establishing productive HIV infection in peripheral blood T lymphocytes and monocyte/macrophages. J Acquir Immune Defic Syndr *7*, 908-915.

Gandhi, S.K., Siliciano, J.D., Bailey, J.R., Siliciano, R.F., and Blankson, J.N. (2008). Role of APOBEC3G/F-mediated hypermutation in the control of human immunodeficiency virus type 1 in elite suppressors. Journal of Virology *82*, 3125-3130.

Garrett, E.D., Tiley, L.S., and Cullen, B.R. (1991). Rev activates expression of the human immunodeficiency virus type 1 vif and vpr gene products. Journal of Virology *65*, 1653-1657.

Gervaix, A., West, D., Leoni, L.M., Richman, D.D., Wong-Staal, F., and Corbeil, J. (1997). A new reporter cell line to monitor HIV infection and drug susceptibility *in vitro*. Proceedings of the National Academy of Sciences of the United States of America *94*, 4653-4658.

Gibbs, J.S., Regier, D.A., and Desrosiers, R.C. (1994). Construction and *in vitro* properties of HIV mutants with deletions in "nonessential" genes. AIDS Research and Human Retroviruses *10*, 343-350.

Gifford, R.J., Katzourakis, A., Tristem, M., Pybus, O.G., Winters, M., and Shafer, R.W. (2008). A transitional endogenous lentivirus from the genome of a basal primate and implications for lentivirus evolution. Proceedings of the National Academy of Sciences of the United States of America *105*, 20362-20367.

Goila-Gaur, R., Khan, M.A., Miyagi, E., Kao, S., and Strebel, K. (2007). Targeting APOBEC3A to the viral nucleoprotein complex confers antiviral activity. Retrovirology *4*, 61.

Goila-Gaur, R., and Strebel, K. (2008). HIV Vif, APOBEC, and intrinsic immunity. Retrovirology *5*, 51.

Goldstone, D.C., Ennis-Adeniran, V., Hedden, J.J., Groom, H.C., Rice, G.I., Christodoulou, E., Walker, P.A., Kelly, G., Haire, L.F., Yap, M.W.*, et al.* (2011). HIV restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. Nature *480*, 379-382.

Goncalves, J., Korin, Y., Zack, J., and Gabuzda, D. (1996). Role of Vif in human immunodeficiency virus type 1 reverse transcription. Journal of Virology *70*, 8701-8709.

Gonda, D.K., Bachmair, A., Wunning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. (1989). Universality and structure of the N-end rule. The Journal of Biological Chemistry *264*, 16700-16712.

Gooch, B.D., and Cullen, B.R. (2008). Functional domain organization of human APOBEC3G. Virology *379*, 118-124.

Gramberg, T., Sunseri, N., and Landau, N.R. (2009). Accessories to the crime: recent advances in HIV accessory protein biology. Curr HIV/AIDS Rep *6*, 36-42.

Guo, F., Cen, S., Niu, M., Saadatmand, J., and Kleiman, L. (2006). Inhibition of formula-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. Journal of Virology *80*, 11710-11722.

Guo, F., Cen, S., Niu, M., Yang, Y., Gorelick, R.J., and Kleiman, L. (2007). The interaction of APOBEC3G with human immunodeficiency virus type 1 nucleocapsid inhibits tRNA3Lys annealing to viral RNA. Journal of Virology *81*, 11322-11331.

Gupta, R.K., Mlcochova, P., Pelchen-Matthews, A., Petit, S.J., Mattiuzzo, G., Pillay, D., Takeuchi, Y., Marsh, M., and Towers, G.J. (2009). Simian immunodeficiency virus envelope glycoprotein counteracts tetherin/BST-2/CD317 by intracellular sequestration. Proceedings of the National Academy of Sciences of the United States of America *106*, 20889-20894.

Haché, G., Abbink, T.E., Berkhout, B., and Harris, R.S. (2009). Optimal translation initiation enables Vif-deficient human immunodeficiency virus type 1 to escape restriction by APOBEC3G. Journal of Virology *83*, 5956-5960.

Haché, G., Liddament, M.T., and Harris, R.S. (2005). The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. The Journal of Biological Chemistry *280*, 10920-10924.

Haché, G., Mansky, L.M., and Harris, R.S. (2006). Human APOBEC3 proteins, retrovirus restriction, and HIV drug resistance. AIDS Rev *8*, 148-157.

Haché, G., Shindo, K., Albin, J.S., and Harris, R.S. (2008). Evolution of HIV isolates that use a novel Vif-independent mechanism to resist restriction by human APOBEC3G. Curr Biol *18*, 819-824.

Hakata, Y., and Landau, N.R. (2006). Reversed functional organization of mouse and human APOBEC3 cytidine deaminase domains. The Journal of Biological Chemistry *281*, 36624-36631.

Harari, A., Ooms, M., Mulder, L.C., and Simon, V. (2009). Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H. Journal of Virology *83*, 295-303.

Harjes, E., Gross, P.J., Chen, K.M., Lu, Y., Shindo, K., Nowarski, R., Gross, J.D., Kotler, M., Harris, R.S., and Matsuo, H. (2009). An extended structure of the APOBEC3G catalytic domain suggests a unique holoenzyme model. J Mol Biol *389*, 819-832.

Harris, R.S. (2008). Enhancing immunity to HIV through APOBEC. Nat Biotechnol *26*, 1089-1090.

Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S., and Malim, M.H. (2003a). DNA deamination mediates innate immunity to retroviral infection. Cell *113*, 803-809.

Harris, R.S., Hultquist, J.F., and Evans, D.T. (2012). The Restriction Factors of Human Immunodeficiency Virus. The Journal of Biological Chemistry *In press*

Harris, R.S., and Liddament, M.T. (2004). Retroviral restriction by APOBEC proteins. Nat Rev Immunol *4*, 868-877.

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

Harris, R.S., Sheehy, A.M., Craig, H.M., Malim, M.H., and Neuberger, M.S. (2003b). DNA deamination: not just a trigger for antibody diversification but also a mechanism for defense against retroviruses. Nat Immunol *4*, 641-643.

Hassaine, G., Agostini, I., Candotti, D., Bessou, G., Caballero, M., Agut, H., Autran, B., Barthalay, Y., and Vigne, R. (2000). Characterization of human immunodeficiency virus type 1 vif gene in long-term asymptomatic individuals. Virology *276*, 169-180.

He, Z., Zhang, W., Chen, G., Xu, R., and Yu, X.F. (2008). Characterization of conserved motifs in HIV Vif required for APOBEC3G and APOBEC3F interaction. J Mol Biol *381*, 1000-1011.

Henriet, S., Mercenne, G., Bernacchi, S., Paillart, J.C., and Marquet, R. (2009). 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 *73*, 211-232.

Holden, L.G., Prochnow, C., Chang, Y.P., Bransteitter, R., Chelico, L., Sen, U., Stevens, R.C., Goodman, M.F., and Chen, X.S. (2008). Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. Nature *456*, 121-124.

Holmes, R.K., Koning, F.A., Bishop, K.N., and Malim, M.H. (2007). APOBEC3F can inhibit the accumulation of HIV reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. The Journal of Biological Chemistry *282*, 2587-2595.

Hrecka, K., Hao, C., Gierszewska, M., Swanson, S.K., Kesik-Brodacka, M., Srivastava, S., Florens, L., Washburn, M.P., and Skowronski, J. (2011). Vpx relieves

inhibition of HIV infection of macrophages mediated by the SAMHD1 protein. Nature *474*, 658-661.

Hultquist, J.F., Binka, M., LaRue, R.S., Simon, V., and Harris, R.S. (2012). Vif Proteins of Human and Simian Immunodeficiency Viruses Require Cellular CBFβ to Degrade APOBEC3 Restriction Factors. Journal of Virology *68*, 2874-2877.

Hultquist, J.F., Lengyel, J.A., Refsland, E.W., Larue, R.S., Lackey, L., Brown, W.L., and Harris, R.S. (2011). Human and Rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H Demonstrate a Conserved Capacity to Restrict Vif-deficient HIV. Journal of Virology *85*, 11220-11234.

Huthoff, H., Autore, F., Gallois-Montbrun, S., Fraternali, F., and Malim, M.H. (2009). RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV. PLoS Pathog *5*, e1000330.

Huthoff, H., and Malim, M.H. (2007). Identification of amino acid residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation. Journal of Virology *81*, 3807-3815.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973-977.

Itaya, S., Nakajima, T., Kaur, G., Terunuma, H., Ohtani, H., Mehra, N., and Kimura, A. (2010). No evidence of an association between the APOBEC3B deletion polymorphism and susceptibility to HIV infection and AIDS in Japanese and Indian populations. J Infect Dis *202*, 815-816; author reply 816-817.

Iwabu, Y., Kinomoto, M., Tatsumi, M., Fujita, H., Shimura, M., Tanaka, Y., Ishizaka, Y., Nolan, D., Mallal, S., Sata, T.*, et al.* (2010). Differential anti-APOBEC3G activity of HIV Vif proteins derived from different subtypes. The Journal of Biological Chemistry *285*, 35350-35358.

Iwatani, Y., Chan, D.S., Liu, L., Yoshii, H., Shibata, J., Yamamoto, N., Levin, J.G., Gronenborn, A.M., and Sugiura, W. (2009). HIV Vif-mediated ubiquitination/degradation of APOBEC3G involves four critical lysine residues in its C-terminal domain. Proceedings of the National Academy of Sciences of the United States of America *106*, 19539-19544.

Iwatani, Y., Chan, D.S., Wang, F., Maynard, K.S., Sugiura, W., Gronenborn, A.M., Rouzina, I., Williams, M.C., Musier-Forsyth, K., and Levin, J.G. (2007). Deaminase-independent inhibition of HIV reverse transcription by APOBEC3G. Nucleic Acids Research *35*, 7096-7108.

Iwatani, Y., Takeuchi, H., Strebel, K., and Levin, J.G. (2006). Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. Journal of Virology *80*, 5992-6002.

Jäger, S., Kim, D.Y., Hultquist, J.F., Shindo, K., Larue, R.S., Kwon, E., Li, M., Anderson, B.D., Yen, L., Stanley, D.*, et al.* (2011). Vif hijacks CBFβ to degrade APOBEC3G and promote HIV infection. Nature *481*, 371-375.

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

Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J., and Navaratnam, N. (2002). An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics *79*, 285-296.

Jern, P., Russell, R.A., Pathak, V.K., and Coffin, J.M. (2009). Likely role of APOBEC3G-mediated G-to-A mutations in HIV evolution and drug resistance. PLoS Pathog *5*, e1000367.

Jia, B., Serra-Moreno, R., Neidermyer, W., Rahmberg, A., Mackey, J., Fofana, I.B., Johnson, W.E., Westmoreland, S., and Evans, D.T. (2009). Species-specific activity of SIV Nef and HIV Vpu in overcoming restriction by tetherin/BST2. PLoS Pathog *5*, e1000429.

Jin, X., Brooks, A., Chen, H., Bennett, R., Reichman, R., and Smith, H. (2005). APOBEC3G/CEM15 (hA3G) mRNA levels associate inversely with human immunodeficiency virus viremia. Journal of Virology *79*, 11513-11516.

Kaiser, S.M., and Emerman, M. (2006). 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. Journal of Virology *80*, 875-882.

Kan, N.C., Franchini, G., Wong-Staal, F., DuBois, G.C., Robey, W.G., Lautenberger, J.A., and Papas, T.S. (1986). Identification of HTLV-III/LAV sor gene product and detection of antibodies in human sera. Science *231*, 1553-1555.

Kao, S., Khan, M.A., Miyagi, E., Plishka, R., Buckler-White, A., and Strebel, K. (2003). The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity. Journal of Virology *77*, 11398-11407.

Kidd, J.M., Newman, T.L., Tuzun, E., Kaul, R., and Eichler, E.E. (2007). Population stratification of a common APOBEC gene deletion polymorphism. PLoS Genet *3*, e63.

Kieffer, T.L., Kwon, P., Nettles, R.E., Han, Y., Ray, S.C., and Siliciano, R.F. (2005). G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells *in vivo*. Journal of Virology *79*, 1975-1980.

Kim, E.Y., Bhattacharya, T., Kunstman, K., Swantek, P., Koning, F.A., Malim, M.H., and Wolinsky, S.M. (2010). Human APOBEC3G-mediated editing can promote HIV sequence diversification and accelerate adaptation to selective pressure. Journal of Virology *84*, 10402-10405.

Kinomoto, M., Kanno, T., Shimura, M., Ishizaka, Y., Kojima, A., Kurata, T., Sata, T., and Tokunaga, K. (2007). All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Research *35*, 2955-2964.

Kitamura, S., Ode, H., Nakashima, M., Imahashi, M., Naganawa, Y., Kurosawa, T., Yokomaku, Y., Yamane, T., Watanabe, N., Suzuki, A.*, et al.* (2012). The APOBEC3C crystal structure and the interface for HIV Vif binding. Nat Struct Mol Biol *19*, 1005-1010.

Klarmann, G.J., Chen, X., North, T.W., and Preston, B.D. (2003). Incorporation of uracil into minus strand DNA affects the specificity of plus strand synthesis initiation

during lentiviral reverse transcription. The Journal of Biological Chemistry *278*, 7902-7909.

Klasse, P.J., and Moore, J.P. (2004). Is there enough gp120 in the body fluids of HIV-infected individuals to have biologically significant effects? Virology *323*, 1-8.

Kleiger, G., Grothe, R., Mallick, P., and Eisenberg, D. (2002). GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. Biochemistry *41*, 5990-5997.

Kohli, R.M., Abrams, S.R., Gajula, K.S., Maul, R.W., Gearhart, P.J., and Stivers, J.T. (2009). A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. The Journal of Biological Chemistry *284*, 22898-22904.

Kohli, R.M., Maul, R.W., Guminski, A.F., McClure, R.L., Gajula, K.S., Saribasak, H., McMahon, M.A., Siliciano, R.F., Gearhart, P.J., and Stivers, J.T. (2010). Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. The Journal of Biological Chemistry *285*, 40956-40964.

Komoto, S., Tsuji, S., Lee, B.J., Iwabu, Y., Kojima, Y., Otake, T., Taniguchi, K., and Ikuta, K. (2005). Higher frequency of premature stop codon mutations at vpu gene of human immunodeficiency virus type 1 CRF01_AE compared with those of other subtypes. Microbes and Infection / Institut Pasteur *7*, 139-147.

Koning, F.A., Goujon, C., Bauby, H., and Malim, M.H. (2011). Target cell-mediated editing of HIV cDNA by APOBEC3 proteins in human macrophages. Journal of Virology *85*, 13448-13452.

Koning, F.A., Newman, E.N., Kim, E.Y., Kunstman, K.J., Wolinsky, S.M., and Malim, M.H. (2009). Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. Journal of Virology *83*, 9474-9485.

Krieger, E., Koraimann, G., and Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. Proteins *47*, 393-402.

Krokan, H.E., Drablos, F., and Slupphaug, G. (2002). Uracil in DNA--occurrence, consequences and repair. Oncogene *21*, 8935-8948.

Laguette, N., Sobhian, B., Casartelli, N., Ringeard, M., Chable-Bessia, C., Segeral, E., Yatim, A., Emiliani, S., Schwartz, O., and Benkirane, M. (2011). SAMHD1 is the dendritic- and myeloid-cell-specific HIV restriction factor counteracted by Vpx. Nature *474*, 654-657.

Land, A.M., Ball, T.B., Luo, M., Pilon, R., Sandstrom, P., Embree, J.E., Wachihi, C., Kimani, J., and Plummer, F.A. (2008). Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya. Journal of Virology *82*, 8172-8182.

Langlois, M.A., Beale, R.C., Conticello, S.G., and Neuberger, M.S. (2005). Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. Nucleic Acids Research *33*, 1913-1923.

Langlois, M.A., and Neuberger, M.S. (2008). Human APOBEC3G can restrict retroviral infection in avian cells and acts independently of both UNG and SMUG1. Journal of Virology *82*, 4660-4664.

LaRue, R.S., Andrésdóttir, V., Blanchard, Y., Conticello, S.G., Derse, D., Emerman, M., Greene, W.C., Jónsson, S.R., Landau, N.R., Löchelt, M.*, et al.* (2009). Guidelines for naming nonprimate APOBEC3 genes and proteins. Journal of Virology *83*, 494-497.

LaRue, R.S., J.A. Lengyel, S.R. Jónsson, V. Andrésdóttir and R.S. Harris (2010). Lentiviral Vif Invariably Degrades the APOBEC3Z3 Protein of Its Mammalian Host but It Is Frequently Capable of Cross-Species Activity. Journal of Virology *84*, 8193-8201.

LaRue, R.S., Jónsson, S.R., Silverstein, K.A.T., Lajoie, M., Bertrand, D., El-Mabrouk, N., Hötzel, I., Andrésdóttir, V., Smith, T.P.L., and Harris, R.S. (2008). 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 *9*, 104 (120 pages).

Lassen, K.G., Wissing, S., Lobritz, M.A., Santiago, M., and Greene, W.C. (2010). Identification of two APOBEC3F splice variants displaying HIV antiviral activity and contrasting sensitivity to Vif. The Journal of Biological Chemistry *285*, 29326-29335.

Lavens, D., Peelman, F., Van der Heyden, J., Uyttendaele, I., Catteeuw, D., Verhee, A., Van Schoubroeck, B., Kurth, J., Hallenberger, S., Clayton, R.*, et al.* (2010). Definition of the interacting interfaces of Apobec3G and HIV Vif using MAPPIT mutagenesis analysis. Nucleic Acids Research *38*, 1902-1912.

Le Tortorec, A., and Neil, S.J. (2009). Antagonism to and intracellular sequestration of human tetherin by the human immunodeficiency virus type 2 envelope glycoprotein. Journal of Virology *83*, 11966-11978.

Lecossier, D., Bouchonnet, F., Clavel, F., and Hance, A.J. (2003). Hypermutation of HIV DNA in the absence of the Vif protein. Science *300*, 1112.

Lee, T.H., Coligan, J.E., Allan, J.S., McLane, M.F., Groopman, J.E., and Essex, M. (1986). A new HTLV-III/LAV protein encoded by a gene found in cytopathic retroviruses. Science *231*, 1546-1549.

Li, J., Potash, M.J., and Volsky, D.J. (2004). Functional domains of APOBEC3G required for antiviral activity. J Cell Biochem *92*, 560-572.

Li, M., Shandilya, S.M., Carpenter, M.A., Rathore, A., Brown, W.L., Perkins, A.L., Harki, D.A., Solberg, J., Hook, D.J., Pandey, K.K.*, et al.* (2012). First-In-Class Small Molecule Inhibitors of the Single-Strand DNA Cytosine Deaminase APOBEC3G. ACS Chem Biol *7*, 506-517.

Li, M.M., Wu, L.I., and Emerman, M. (2010). The range of human APOBEC3H sensitivity to lentiviral Vif proteins. Journal of Virology *84*, 88-95.

Li, Q., Smith, A.J., Schacker, T.W., Carlis, J.V., Duan, L., Reilly, C.S., and Haase, A.T. (2009). Microarray analysis of lymphatic tissue reveals stage-specific, gene expression signatures in HIV infection. J Immunol *183*, 1975-1982.

331

Li, X., Ma, J., Zhang, Q., Zhou, J., Yin, X., Zhai, C., You, X., Yu, L., Guo, F., Zhao, L.*, et al.* (2011). Functional analysis of the two cytidine deaminase domains in APOBEC3G. Virology *414*, 130-136.

Li, X.Y., Guo, F., Zhang, L., Kleiman, L., and Cen, S. (2007). APOBEC3G inhibits DNA strand transfer during HIV reverse transcription. The Journal of Biological Chemistry *282*, 32065-32074.

Liddament, M.T., Brown, W.L., Schumacher, A.J., and Harris, R.S. (2004). APOBEC3F properties and hypermutation preferences indicate activity against HIV *in vivo*. Curr Biol *14*, 1385-1391.

Lim, E.S., Fregoso, O.I., McCoy, C.O., Matsen, F.A., Malik, H.S., and Emerman, M. (2012). The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. Cell Host Microbe *11*, 194-204.

Liu, J., and Nussinov, R. (2009). The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: conformational control of substrate orientation. PLoS Comput Biol *5*, e1000527.

Liu, J., and Nussinov, R. (2010). Molecular dynamics reveal the essential role of linker motions in the function of cullin-RING E3 ligases. J Mol Biol *396*, 1508-1523.

Liu, J., and Nussinov, R. (2011). Flexible cullins in cullin-RING E3 ligases allosterically regulate ubiquitination. The Journal of Biological Chemistry *286*, 40934-40942.

Löchelt, M., Romen, F., Bastone, P., Muckenfuss, H., Kirchner, N., Kim, Y.B., Truyen, U., Rösler, U., Battenberg, M., Saib, A.*, et al.* (2005). The antiretroviral activity of APOBEC3 is inhibited by the foamy virus accessory Bet protein. Proceedings of the National Academy of Sciences of the United States of America *102*, 7982-7987.

Loeb, L.A., Essigmann, J.M., Kazazi, F., Zhang, J., Rose, K.D., and Mullins, J.I. (1999). Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proceedings of the National Academy of Sciences of the United States of America *96*, 1492-1497.

Lorieau, J.L., Louis, J.M., and Bax, A. (2010). The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. Proceedings of the National Academy of Sciences of the United States of America *107*, 11341-11346.

Luo, K., Liu, B., Xiao, Z., Yu, Y., Yu, X., Gorelick, R., and Yu, X.F. (2004). Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging. Journal of Virology *78*, 11841-11852.

Luo, K., Wang, T., Liu, B., Tian, C., Xiao, Z., Kappes, J., and Yu, X.F. (2007). Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation. Journal of Virology *81*, 7238-7248.

Luo, K., Xiao, Z., Ehrlich, E., Yu, Y., Liu, B., Zheng, S., and Yu, X.F. (2005). Primate lentiviral virion infectivity factors are substrate receptors that assemble with cullin 5-E3 ligase through a HCCH motif to suppress APOBEC3G. Proceedings of the

National Academy of Sciences of the United States of America *102*, 11444-11449.

Madani, N., and Kabat, D. (1998). An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. Journal of Virology *72*, 10251-10255.

Malim, M.H. (2009). APOBEC proteins and intrinsic resistance to HIV infection. Philos Trans R Soc Lond B Biol Sci *364*, 675-687.

Malim, M.H., and Bieniasz, P.D. (2012). HIV Restriction Factors and Mechanisms of Evasion. Cold Spring Harb Perspect Med *2*, a006940.

Malim, M.H., and Emerman, M. (2008). HIV accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe *3*, 388-398.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature *424*, 99-103.

Mangeat, B., Turelli, P., Liao, S., and Trono, D. (2004). A single amino acid determinant governs the species-specific sensitivity of APOBEC3G to Vif action. The Journal of Biological Chemistry *279*, 14481-14483.

Mansky, L.M., Preveral, S., Selig, L., Benarous, R., and Benichou, S. (2000). The interaction of vpr with uracil DNA glycosylase modulates the human immunodeficiency virus type 1 *In vivo* mutation rate. Journal of Virology *74*, 7039-7047.

Mansky, L.M., and Temin, H.M. (1995). Lower *in vivo* mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. Journal of Virology *69*, 5087-5094.

Mariani, R., Chen, D., Schröfelbauer, B., Navarro, F., König, R., Bollman, B., Münk, C., Nymark-McMahon, H., and Landau, N.R. (2003). Species-specific exclusion of APOBEC3G from HIV virions by Vif. Cell *114*, 21-31.

Marin, M., Golem, S., Rose, K.M., Kozak, S.L., and Kabat, D. (2008). Human immunodeficiency virus type 1 Vif functionally interacts with diverse APOBEC3 cytidine deaminases and moves with them between cytoplasmic sites of mRNA metabolism. Journal of Virology *82*, 987-998.

Marin, M., Rose, K.M., Kozak, S.L., and Kabat, D. (2003). HIV Vif protein binds the editing enzyme APOBEC3G and induces its degradation. Nat Med *9*, 1398-1403.

Mbisa, J.L., Barr, R., Thomas, J.A., Vandegraaff, N., Dorweiler, I.J., Svarovskaia, E.S., Brown, W.L., Mansky, L.M., Gorelick, R.J., Harris, R.S.*, et al.* (2007). HIV cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. Journal of Virology *81*, 7099-7110.

Mbisa, J.L., Bu, W., and Pathak, V.K. (2010). APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms. Journal of Virology *84*, 5250-5259.

Mehle, A., Goncalves, J., Santa-Marta, M., McPike, M., and Gabuzda, D. (2004a). Phosphorylation of a novel SOCS-box regulates assembly of the HIV Vif-Cul5 complex that promotes APOBEC3G degradation. Genes Dev *18*, 2861-2866.

Mehle, A., Strack, B., Ancuta, P., Zhang, C., McPike, M., and Gabuzda, D. (2004b). Vif overcomes the innate antiviral activity of APOBEC3G by promoting its

degradation in the ubiquitin-proteasome pathway. The Journal of Biological Chemistry *279*, 7792-7798.

Mehle, A., Thomas, E.R., Rajendran, K.S., and Gabuzda, D. (2006). A zinc-binding region in Vif binds Cul5 and determines cullin selection. The Journal of Biological Chemistry *281*, 17259-17265.

Mercenne, G., Bernacchi, S., Richer, D., Bec, G., Henriet, S., Paillart, J.C., and Marquet, R. (2009). HIV Vif binds to APOBEC3G mRNA and inhibits its translation. Nucleic Acids Research *38*, 633-646.

Miyagi, E., Brown, C.R., Opi, S., Khan, M., Goila-Gaur, R., Kao, S., Walker, R.C., Jr., Hirsch, V., and Strebel, K. (2010). Stably expressed APOBEC3F has negligible antiviral activity. Journal of Virology *84*, 11067-11075.

Miyagi, E., Opi, S., Takeuchi, H., Khan, M., Goila-Gaur, R., Kao, S., and Strebel, K. (2007). Enzymatically active APOBEC3G is required for efficient inhibition of human immunodeficiency virus type 1. Journal of Virology *81*, 13346-13353.

Muckenfuss, H., Hamdorf, M., Held, U., Perkovic, M., Lower, J., Cichutek, K., Flory, E., Schumann, G.G., and Munk, C. (2006). APOBEC3 proteins inhibit human LINE-1 retrotransposition. The Journal of Biological Chemistry *281*, 22161-22172.

Mulder, L.C., Harari, A., and Simon, V. (2008). Cytidine deamination induced HIV drug resistance. Proceedings of the National Academy of Sciences of the United States of America *105*, 5501-5506.

Mulder, L.C., Ooms, M., Majdak, S., Smedresman, J., Linscheid, C., Harari, A., Kunz, A., and Simon, V. (2010). Moderate influence of human APOBEC3F on HIV replication in primary lymphocytes. Journal of Virology *84*, 9613-9617.

Nagao, T., Yamashita, T., Miyake, A., Uchiyama, T., Nomaguchi, M., and Adachi, A. (2010). Different interaction between HIV Vif and its cellular target proteins APOBEC3G/APOBEC3F. J Med Invest *57*, 89-94.

Nakaya, T., Fujinaga, K., Kishi, M., Oka, S., Kurata, T., Jones, I.M., and Ikuta, K. (1994). Nonsense mutations in the vpr gene of HIV during *in vitro* virus passage and in HIV carrier-derived peripheral blood mononuclear cells. FEBS Lett *354*, 17-22.

Nathans, R., Cao, H., Sharova, N., Ali, A., Sharkey, M., Stranska, R., Stevenson, M., and Rana, T.M. (2008). Small-molecule inhibition of HIV Vif. Nat Biotechnol *26*, 1187-1192.

Navarro, F., Bollman, B., Chen, H., Konig, R., Yu, Q., Chiles, K., and Landau, N.R. (2005). Complementary function of the two catalytic domains of APOBEC3G. Virology *333*, 374-386.

Neil, S.J., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV Vpu. Nature *451*, 425-430.

Newman, E.N., Holmes, R.K., Craig, H.M., Klein, K.C., Lingappa, J.R., Malim, M.H., and Sheehy, A.M. (2005). Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. Curr Biol *15*, 166-170.

Nowarski, R., Britan-Rosich, E., Shiloach, T., and Kotler, M. (2008). Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase. Nat Struct Mol Biol *15*, 1059-1066.

OhAinle, M., Kerns, J.A., Li, M.M., Malik, H.S., and Emerman, M. (2008). Antiretroelement activity of APOBEC3H was lost twice in recent human evolution. Cell Host Microbe *4*, 249-259.

OhAinle, M., Kerns, J.A., Malik, H.S., and Emerman, M. (2006). Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H. Journal of Virology *80*, 3853-3862.

Okeoma, C.M., Lovsin, N., Peterlin, B.M., and Ross, S.R. (2007). APOBEC3 inhibits mouse mammary tumour virus replication *in vivo*. Nature *445*, 927-930.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Opi, S., Takeuchi, H., Kao, S., Khan, M.A., Miyagi, E., Goila-Gaur, R., Iwatani, Y., Levin, J.G., and Strebel, K. (2006). Monomeric APOBEC3G is catalytically active and has antiviral activity. Journal of Virology *80*, 4673-4682.

Pace, C., Keller, J., Nolan, D., James, I., Gaudieri, S., Moore, C., and Mallal, S. (2006). Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. Journal of Virology *80*, 9259-9269.

Pancera, M., Majeed, S., Ban, Y.E., Chen, L., Huang, C.C., Kong, L., Kwon, Y.D., Stuckey, J., Zhou, T., Robinson, J.E.*, et al.* (2010). Structure of HIV gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. Proceedings of the National Academy of Sciences of the United States of America *107*, 1166-1171.

Pathak, V.K., and Temin, H.M. (1990). 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. Proceedings of the National Academy of Sciences of the United States of America *87*, 6019-6023.

Peng, G., Greenwell-Wild, T., Nares, S., Jin, W., Lei, K.J., Rangel, Z.G., Munson, P.J., and Wahl, S.M. (2007). Myeloid differentiation and susceptibility to HIV are linked to APOBEC3 expression. Blood *110*, 393-400.

Peng, G., Lei, K.J., Jin, W., Greenwell-Wild, T., and Wahl, S.M. (2006). Induction of APOBEC3 family proteins, a defensive maneuver underlying interferon-induced anti-HIV activity. J Exp Med *203*, 41-46.

Pery, E., Rajendran, K.S., Brazier, A.J., and Gabuzda, D. (2009). Regulation of APOBEC3 proteins by a novel YXXL motif in human immunodeficiency virus type 1 Vif and simian immunodeficiency virus SIVagm Vif. Journal of Virology *83*, 2374-2381.

Pertel, T., Hausmann, S., Morger, D., Züger, S., Guerra, J., Lascano, J., Reinhard, C., Santoni, F.A., Uchill, P.D., Chatel, L., *et al.* (2011). TRIM5 is an innate immune sensor for the retrovirus capsid lattice. Nature *472*, 361-365.

Petersen-Mahrt, S.K., Harris, R.S., and Neuberger, M.S. (2002). AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. Nature *418*, 99-103.

Piantadosi, A., Humes, D., Chohan, B., McClelland, R.S., and Overbaugh, J. (2009). 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. Journal of Virology *83*, 7805-7814.

Pillai, S.K., Abdel-Mohsen, M., Guatelli, J., Skasko, M., Monto, A., Fujimoto, K., Yukl, S., Greene, W.C., Kovari, H., Rauch, A.*, et al.* (2012). Role of retroviral restriction factors in the interferon-alpha-mediated suppression of HIV *in vivo*. Proceedings of the National Academy of Sciences of the United States of America *109*, 3035-3040.

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

Pion, M., Granelli-Piperno, A., Mangeat, B., Stalder, R., Correa, R., Steinman, R.M., and Piguet, V. (2006). APOBEC3G/3F mediates intrinsic resistance of monocyte-derived dendritic cells to HIV infection. J Exp Med *203*, 2887-2893.

Powell, R.D., Holland, P.J., Hollis, T., and Perrino, F.W. (2011). Aicardi-Goutières Syndrome Gene and HIV Restriction Factor SAMHD1 Is a dGTP-regulated Deoxynucleotide Triphosphohydrolase. The Journal of Biological Chemistry *286*, 43596-43600.

Priet, S., Gros, N., Navarro, J.M., Boretto, J., Canard, B., Querat, G., and Sire, J. (2005). HIV-associated uracil DNA glycosylase activity controls dUTP misincorporation in viral DNA and is essential to the HIV life cycle. Mol Cell *17*, 479-490.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201-1212.

Rangel, H.R., Garzaro, D., Rodriguez, A.K., Ramirez, A.H., Ameli, G., Del Rosario Gutierrez, C., and Pujol, F.H. (2009). Deletion, insertion and stop codon mutations in vif genes of HIV infecting slow progressor patients. Journal of infection in developing countries *3*, 531-538.

Rasband, W.S. (1997-2009). ImageJ. National Institutes of Health, Bethesda, Maryland, USA *http://rsb.info.nih.gov/ij/.*

Rausch, J.W., Chelico, L., Goodman, M.F., and LeGrice, S.F. (2009). Dissecting APOBEC3G substrate specificity by nucleoside analog interference. The Journal of biological chemistry *284*, 7047-7058.

Reddy, K., Winkler, C.A., Werner, L., Mlisana, K., Abdool Karim, S.S., and Ndung'u, T. (2010). APOBEC3G expression is dysregulated in primary HIV infection and polymorphic variants influence CD4+ T-cell counts and plasma viral load. AIDS *24*, 195-204.

Refsland, E.W., Hultquist, J.F., and Harris, R.S. (2012). Endogenous origins of HIV G-to-A hypermutation. PLoS Pathog *8*, e1002800.

Refsland, E.W., Stenglein, M.D., Shindo, K., Albin, J.S., Brown, W.L., and Harris, R.S. (2010). Quantitative profiling of the full APOBEC3 mRNA repertoire in

lymphocytes and tissues: implications for HIV restriction. Nucleic Acids Research *38*, 4274-4284.

Reingewertz, T.H., Shalev, D.E., and Friedler, A. (2010). Structural disorder in the HIV-1 Vif protein and interaction-dependent gain of structure. Protein Pept Lett *17*, 988-998.

Rose, K.M., Marin, M., Kozak, S.L., and Kabat, D. (2005). Regulated production and anti-HIV type 1 activities of cytidine deaminases APOBEC3B, 3F, and 3G. AIDS Research and Human Retroviruses *21*, 611-619.

Russell, R.A., Moore, M.D., Hu, W.S., and Pathak, V.K. (2009a). APOBEC3G induces a hypermutation gradient: purifying selection at multiple steps during HIV replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. Retrovirology *6*, 16.

Russell, R.A., and Pathak, V.K. (2007). Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F. Journal of Virology *81*, 8201-8210.

Russell, R.A., Smith, J., Barr, R., Bhattacharyya, D., and Pathak, V.K. (2009b). Distinct domains within APOBEC3G and APOBEC3F interact with separate regions of human immunodeficiency virus type 1 Vif. Journal of Virology *83*, 1992-2003.

Russell, R.A., Wiegand, H.L., Moore, M.D., Schafer, A., McClure, M.O., and Cullen, B.R. (2005). Foamy virus Bet proteins function as novel inhibitors of the APOBEC3 family of innate antiretroviral defense factors. Journal of Virology *79*, 8724-8731.

Sadler, H.A., Stenglein, M.D., Harris, R.S., and Mansky, L.M. (2010). APOBEC3G contributes to HIV variation through sublethal mutagenesis. Journal of Virology *84*, 7396-7404.

Santa-Marta, M., da Silva, F.A., Fonseca, A.M., and Goncalves, J. (2005). HIV 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. The Journal of Biological Chemistry *280*, 8765-8775.

Santosuosso, M., Righi, E., Lindstrom, V., Leblanc, P.R., and Poznansky, M.C. (2009). HIV envelope protein gp120 is present at high concentrations in secondary lymphoid organs of individuals with chronic HIV infection. J Infect Dis *200*, 1050-1053.

Sauter, D., Schindler, M., Specht, A., Landford, W.N., Munch, J., Kim, K.A., Votteler, J., Schubert, U., Bibollet-Ruche, F., Keele, B.F.*, et al.* (2009). Tetherin-driven adaptation of Vpu and Nef function and the evolution of pandemic and nonpandemic HIV strains. Cell Host Microbe *6*, 409-421.

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

Sayah, D.M., Sokolskaja, E., Berthoux, L., and Luban, J. (2004). Cyclophilin A retrotransposition into TRIM5 explains owl monkey resistance to HIV. Nature *430*, 569-573.

Schafer, A., Bogerd, H.P., and Cullen, B.R. (2004). Specific packaging of APOBEC3G into HIV virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. Virology *328*, 163-168.

Scherer, D.C., Brockman, J.A., Chen, Z., Maniatis, T., and Ballard, D.W. (1995). Signal-induced degradation of I kappa B alpha requires site-specific ubiquitination. Proceedings of the National Academy of Sciences of the United States of America *92*, 11259-11263.

Schmitt, K., Hill, M.S., Liu, Z., Ruiz, A., Culley, N., Pinson, D.M., and Stephens, E.B. (2010). Comparison of the replication and persistence of simian-human immunodeficiency viruses expressing Vif proteins with mutation of the SLQYLA or HCCH domains in macaques. Virology *404*, 187-203.

Schmitt, K., Hill, M.S., Ruiz, A., Culley, N., Pinson, D.M., Wong, S.W., and Stephens, E.B. (2009). Mutations in the highly conserved SLQYLA motif of Vif in a simian-human immunodeficiency virus result in a less pathogenic virus and are associated with G-to-A mutations in the viral genome. Virology *383*, 362-372.

Schröfelbauer, B., Chen, D., and Landau, N.R. (2004). A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif). Proceedings of the National Academy of Sciences of the United States of America *101*, 3927-3932.

Schröfelbauer, B., Senger, T., Manning, G., and Landau, N.R. (2006). Mutational alteration of human immunodeficiency virus type 1 Vif allows for functional interaction with nonhuman primate APOBEC3G. Journal of Virology *80*, 5984-5991.

Schröfelbauer, B., Yu, Q., Zeitlin, S.G., and Landau, N.R. (2005). Human immunodeficiency virus type 1 Vpr induces the degradation of the UNG and SMUG uracil-DNA glycosylases. Journal of Virology *79*, 10978-10987.

Schumacher, A.J., Haché, G., MacDuff, D.A., Brown, W.L., and Harris, R.S. (2008). The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. Journal of Virology *82*, 2652-2660.

Schwartz, S., Felber, B.K., and Pavlakis, G.N. (1991). Expression of human immunodeficiency virus type 1 vif and vpr mRNAs is Rev-dependent and regulated by splicing. Virology *183*, 677-686.

Serra-Moreno, R., Jia, B., Breed, M., Alvarez, X., and Evans, D.T. (2011). Compensatory changes in the cytoplasmic tail of gp41 confer resistance to tetherin/BST-2 in a pathogenic nef-deleted SIV. Cell Host Microbe *9*, 46-57.

Shandilya, S.M., Nalam, M.N., Nalivaika, E.A., Gross, P.J., Valesano, J.C., Shindo, K., Li, M., Munson, M., Royer, W.E., Harjes, E.*, et al.* (2010). Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. Structure *18*, 28-38.

Shao, Q., Wang, Y., Hildreth, J.E., and Liu, B. (2010). Polyubiquitination of APOBEC3G is essential for its degradation by HIV Vif. Journal of Virology *84*, 4840-4844.

Sheehy, A.M., and Erthal, J. (2012). APOBEC3 versus Retroviruses, Immunity versus Invasion: Clash of the Titans. Mol Biol Int *2012*, 974924.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV infection and is suppressed by the viral Vif protein. Nature *418*, 646-650.

Sheehy, A.M., Gaddis, N.C., and Malim, M.H. (2003). The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV Vif. Nat Med *9*, 1404-1407.

Shindo, K., Takaori-Kondo, A., Kobayashi, M., Abudu, A., Fukunaga, K., and Uchiyama, T. (2003). The enzymatic activity of CEM15/Apobec-3G is essential for the regulation of the infectivity of HIV virion but not a sole determinant of its antiviral activity. The Journal of Biological Chemistry *278*, 44412-44416.

Shlyakhtenko, L.S., Lushnikov, A.Y., Li, M., Lackey, L., Harris, R.S., and Lyubchenko, Y.L. (2011). Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. The Journal of Biological Chemistry *286*, 3387-3395.

Simon, J.H., Gaddis, N.C., Fouchier, R.A., and Malim, M.H. (1998). Evidence for a newly discovered cellular anti-HIV phenotype. Nat Med *4*, 1397-1400.

Simon, J.H., and Malim, M.H. (1996). The human immunodeficiency virus type 1 Vif protein modulates the postpenetration stability of viral nucleoprotein complexes. Journal of Virology *70*, 5297-5305.

Simon, J.H., Sheehy, A.M., Carpenter, E.A., Fouchier, R.A., and Malim, M.H. (1999). Mutational analysis of the human immunodeficiency virus type 1 Vif protein. Journal of Virology *73*, 2675-2681.

Simon, V., Zennou, V., Murray, D., Huang, Y., Ho, D.D., and Bieniasz, P.D. (2005). Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV diversification. PLoS Pathog *1*, e6.

Smith, J.L., Bu, W., Burdick, R.C., and Pathak, V.K. (2009). Multiple ways of targeting APOBEC3-virion infectivity factor interactions for anti-HIV drug development. Trends Pharmacol Sci *30*, 638-646.

Smith, J.L., and Pathak, V.K. (2010). Identification of specific determinants of human APOBEC3F, APOBEC3C, and APOBEC3DE and African green monkey APOBEC3F that interact with HIV Vif. Journal of Virology *84*, 12599-12608.

Sodroski, J., Goh, W.C., Rosen, C., Tartar, A., Portetelle, D., Burny, A., and Haseltine, W. (1986). Replicative and cytopathic potential of HTLV-III/LAV with sor gene deletions. Science *231*, 1549-1553.

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

Sova, P., van Ranst, M., Gupta, P., Balachandran, R., Chao, W., Itescu, S., McKinley, G., and Volsky, D.J. (1995). Conservation of an intact human immunodeficiency virus type 1 vif gene *in vitro* and *in vivo*. Journal of Virology *69*, 2557-2564.

Sova, P., and Volsky, D.J. (1993). Efficiency of viral DNA synthesis during infection of permissive and nonpermissive cells with vif-negative human immunodeficiency virus type 1. Journal of Virology *67*, 6322-6326.

Stanley, B.J., Ehrlich, E.S., Short, L., Yu, Y., Xiao, Z., Yu, X.F., and Xiong, Y. (2008). Structural Insight into the HIV Vif SOCS Box and Its Role in Human E3 Ubiquitin Ligase Assembly. Journal of Virology *82*, 8656-8663.

Stenglein, M.D., Burns, M.B., Li, M., Lengyel, J., and Harris, R.S. (2010). APOBEC3 proteins mediate the clearance of foreign DNA from human cells. Nat Struct Mol Biol *17*, 222-229.

Stenglein, M.D., and Harris, R.S. (2006). APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. The Journal of Biological Chemistry *281*, 16837-16841.

Stopak, K., de Noronha, C., Yonemoto, W., and Greene, W.C. (2003). HIV Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability. Mol Cell *12*, 591-601.

Stoye, J.P. (2012). Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nat Rev Microbiol *10*, 395-406.

Strebel, K., Daugherty, D., Clouse, K., Cohen, D., Folks, T., and Martin, M.A. (1987). The HIV 'A' (sor) gene product is essential for virus infectivity. Nature *328*, 728-730.

Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV infection in Old World monkeys. Nature *427*, 848-853.

Suspène, R., Rusniok, C., Vartanian, J.P., and Wain-Hobson, S. (2006). Twin gradients in APOBEC3 edited HIV DNA reflect the dynamics of lentiviral replication. Nucleic Acids Research *34*, 4677-4684.

Suspène, R., Sommer, P., Henry, M., Ferris, S., Guetard, D., Pochet, S., Chester, A., Navaratnam, N., Wain-Hobson, S., and Vartanian, J.P. (2004). APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase. Nucleic Acids Research *32*, 2421-2429.

Svarovskaia, E.S., Xu, H., Mbisa, J.L., Barr, R., Gorelick, R.J., Ono, A., Freed, E.O., Hu, W.S., and Pathak, V.K. (2004). Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV virions through interactions with viral and nonviral RNAs. The Journal of Biological Chemistry *279*, 35822-35828.

Tan, L., Sarkis, P.T., Wang, T., Tian, C., and Yu, X.F. (2009). Sole copy of Z2-type human cytidine deaminase APOBEC3H has inhibitory activity against retrotransposons and HIV. FASEB J *23*, 279-287.

Thielen, B.K., Klein, K.C., Walker, L.W., Rieck, M., Buckner, J.H., Tomblingson, G.W., and Lingappa, J.R. (2007). T cells contain an RNase-insensitive inhibitor of APOBEC3G deaminase activity. PLoS Pathog *3*, 1320-1334.

Thielen, B.K., McNevin, J.P., McElrath, M.J., Hunt, B.V., Klein, K.C., and Lingappa, J.R. (2010). Innate immune signaling induces high levels of TC-specific deaminase activity in primary monocyte-derived cells through expression of APOBEC3A isoforms. The Journal of Biological Chemistry *285*, 27753-27766.

Tian, C., Wang, T., Zhang, W., and Yu, X.F. (2007). Virion packaging determinants and reverse transcription of SRP RNA in HIV particles. Nucleic Acids Research *35*, 7288-7302.

Tian, C., Yu, X., Zhang, W., Wang, T., Xu, R., and Yu, X.F. (2006). Differential requirement for conserved tryptophans in human immunodeficiency virus type 1 Vif for the selective suppression of APOBEC3G and APOBEC3F. Journal of Virology *80*, 3112-3115.

Tominaga, K., Kato, S., Negishi, M., and Takano, T. (1996). A high frequency of defective vif genes in peripheral blood mononuclear cells from HIV type 1-infected individuals. AIDS Research and Human Retroviruses *12*, 1543-1549.

Towers, G.J., Hatziioannou, T., Cowan, S., Goff, S.P., Luban, J., and Bieniasz, P.D. (2003). Cyclophilin A modulates the sensitivity of HIV to host restriction factors. Nat Med *9*, 1138-1143.

Turelli, P., Mangeat, B., Jost, S., Vianin, S., and Trono, D. (2004). Inhibition of hepatitis B virus replication by APOBEC3G. Science *303*, 1829.

Ulenga, N.K., Sarr, A.D., Hamel, D., Sankale, J.L., Mboup, S., and Kanki, P.J. (2008a). The level of APOBEC3G (hA3G)-related G-to-A mutations does not correlate with viral load in HIV type 1-infected individuals. AIDS Research and Human Retroviruses *24*, 1285-1290.

Ulenga, N.K., Sarr, A.D., Thakore-Meloni, S., Sankale, J.L., Eisen, G., and Kanki, P.J. (2008b). Relationship between human immunodeficiency type 1 infection and expression of human APOBEC3G and APOBEC3F. J Infect Dis *198*, 486-492.

Valcke, H.S., Bernard, N.F., Bruneau, J., Alary, M., Tsoukas, C.M., and Roger, M. (2006). APOBEC3G genetic variants and their association with risk of HIV infection in highly exposed Caucasians. AIDS *20*, 1984-1986.

Van Damme, N., Goff, D., Katsura, C., Jorgenson, R.L., Mitchell, R., Johnson, M.C., Stephens, E.B., and Guatelli, J. (2008). The interferon-induced protein BST-2 restricts HIV release and is downregulated from the cell surface by the viral Vpu protein. Cell Host Microbe *3*, 245-252.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci.

Vartanian, J.P., Meyerhans, A., Asjo, B., and Wain-Hobson, S. (1991). Selection, recombination, and G-->A hypermutation of human immunodeficiency virus type 1 genomes. Journal of Virology *65*, 1779-1788.

Vartanian, J.P., Meyerhans, A., Sala, M., and Wain-Hobson, S. (1994). G-->A hypermutation of the human immunodeficiency virus type 1 genome: evidence for dCTP pool imbalance during reverse transcription. Proceedings of the National Academy of Sciences of the United States of America *91*, 3092-3096.

Vazquez-Perez, J.A., Ormsby, C.E., Hernandez-Juan, R., Torres, K.J., and Reyes-Teran, G. (2009). APOBEC3G mRNA expression in exposed seronegative and early stage HIV infected individuals decreases with removal of exposure and with disease progression. Retrovirology *6*, 23.

Vetter, M.L., Johnson, M.E., Antons, A.K., Unutmaz, D., and D'Aquila, R.T. (2009). Differences in APOBEC3G expression in CD4+ T helper lymphocyte subtypes modulate HIV infectivity. PLoS Pathog *5*, e1000292.

Virgen, C.A., and Hatziioannou, T. (2007). Antiretroviral activity and Vif sensitivity of rhesus macaque APOBEC3 proteins. Journal of Virology *81*, 13932-13937.

von Schwedler, U., Song, J., Aiken, C., and Trono, D. (1993). Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells. Journal of Virology *67*, 4945-4955.

Wang, T., Tian, C., Zhang, W., Luo, K., Sarkis, P.T., Yu, L., Liu, B., Yu, Y., and Yu, X.F. (2007). 7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G. Journal of Virology *81*, 13112-13124.

Wang, T., Tian, C., Zhang, W., Sarkis, P.T., and Yu, X.F. (2008). Interaction with 7SL RNA but not with HIV genomic RNA or P bodies is required for APOBEC3F virion packaging. J Mol Biol *375*, 1098-1112.

Wang, X., Abudu, A., Son, S., Dang, Y., Venta, P.J., and Zheng, Y.H. (2011a). Analysis of human APOBEC3H haplotypes and anti-human immunodeficiency virus type 1 activity. Journal of Virology *85*, 3142-3152.

Wang, X., Ao, Z., Chen, L., Kobinger, G., Peng, J., and Yao, X. (2012). The cellular antiviral protein APOBEC3G interacts with HIV reverse transcriptase and inhibits its function during viral replication. Journal of Virology *86*, 3777-3786.

Wang, Y., Shao, Q., Yu, X., Kong, W., Hildreth, J.E., and Liu, B. (2011b). N-terminal hemagglutinin tag renders lysine-deficient APOBEC3G resistant to HIV Vif-induced degradation by reduced polyubiquitination. Journal of Virology *85*, 4510-4519.

Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, J.W., Jr., Swanstrom, R., Burch, C.L., and Weeks, K.M. (2009). Architecture and secondary structure of an entire HIV RNA genome. Nature *460*, 711-716.

Wedekind, J.E., Dance, G.S., Sowden, M.P., and Smith, H.C. (2003). Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. Trends Genet *19*, 207-216.

Wedekind, J.E., Gillilan, R., Janda, A., Krucinska, J., Salter, J.D., Bennett, R.P., Raina, J., and Smith, H.C. (2006). Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits. The Journal of Biological Chemistry *281*, 38122-38126.

Wiegand, H.L., Doehle, B.P., Bogerd, H.P., and Cullen, B.R. (2004). A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV and HIV-2 Vif proteins. EMBO J *23*, 2451-2458.

Wieland, U., Hartmann, J., Suhr, H., Salzberger, B., Eggers, H.J., and Kuhn, J.E. (1994). *In vivo* genetic variability of the HIV vif gene. Virology *203*, 43-51.

Wieland, U., Seelhoff, A., Hofmann, A., Kuhn, J.E., Eggers, H.J., Mugyenyi, P., and Schwander, S. (1997). Diversity of the vif gene of human immunodeficiency virus type 1 in Uganda. The Journal of General Virology *78 (Pt 2)*, 393-400.

Wood, N., Bhattacharya, T., Keele, B.F., Giorgi, E., Liu, M., Gaschen, B., Daniels, M., Ferrari, G., Haynes, B.F., McMichael, A.*, et al.* (2009). HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC. PLoS Pathog *5*, e1000414.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E.*, et al.* (2008). Direct evidence of extensive diversity of HIV in Kinshasa by 1960. Nature *455*, 661-664.

Wu, G., Xu, G., Schulman, B.A., Jeffrey, P.D., Harper, J.W., and Pavletich, N.P. (2003). Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. Mol Cell *11*, 1445-1456.

Xiao, Z., Xiong, Y., Zhang, W., Tan, L., Ehrlich, E., Guo, D., and Yu, X.F. (2007). Characterization of a novel Cullin5 binding domain in HIV Vif. J Mol Biol *373*, 541-550.

Xu, G., Paige, J.S., and Jaffrey, S.R. (2010). Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. Nat Biotechnol *28*, 868-873.

Xu, H., Chertova, E., Chen, J., Ott, D.E., Roser, J.D., Hu, W.S., and Pathak, V.K. (2007). Stoichiometry of the antiviral protein APOBEC3G in HIV virions. Virology *360*, 247-256.

Xu, H., Svarovskaia, E.S., Barr, R., Zhang, Y., Khan, M.A., Strebel, K., and Pathak, V.K. (2004). A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV virion infectivity factor-induced depletion. Proceedings of the National Academy of Sciences of the United States of America *101*, 5652-5657.

Yamashita, T., Kamada, K., Hatcho, K., Adachi, A., and Nomaguchi, M. (2008). Identification of amino acid residues in HIV Vif critical for binding and exclusion of APOBEC3G/F. Microbes and Infection / Institut Pasteur *10*, 1142-1149.

Yamashita, T., Nomaguchi, M., Miyake, A., Uchiyama, T., and Adachi, A. (2010). Status of APOBEC3G/F in cells and progeny virions modulated by Vif determines HIV-1 infectivity. Microbes and Infection / Institut Pasteur *12*, 166-171.

Yan, N., O'Day, E., Wheeler, L.A., Engelman, A., and Lieberman, J. (2011). HIV DNA is heavily uracilated, which protects it from autointegration. Proceedings of the National Academy of Sciences of the United States of America *108*, 9244-9249.

Yang, B., Chen, K., Zhang, C., Huang, S., and Zhang, H. (2007a). Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV DNA. The Journal of Biological Chemistry *282*, 11667-11675.

Yang, B., Gao, L., Li, L., Lu, Z., Fan, X., Patel, C.A., Pomerantz, R.J., DuBois, G.C., and Zhang, H. (2003). Potent suppression of viral infectivity by the peptides that inhibit multimerization of human immunodeficiency virus type 1 (HIV) Vif proteins. The Journal of Biological Chemistry *278*, 6596-6602.

Yang, S., Sun, Y., and Zhang, H. (2001). The multimerization of human immunodeficiency virus type I Vif protein: a requirement for Vif function in the viral life cycle. The Journal of Biological Chemistry *276*, 4889-4893.

Yang, Y., Guo, F., Cen, S., and Kleiman, L. (2007b). Inhibition of initiation of reverse transcription in HIV by human APOBEC3F. Virology *365*, 92-100.

Yedavalli, V.R., and Ahmad, N. (2001). Low conservation of functional domains of HIV type 1 vif and vpr genes in infected mothers correlates with lack of vertical transmission. AIDS Research and Human Retroviruses *17*, 911-923.

Yedavalli, V.R., Chappey, C., Matala, E., and Ahmad, N. (1998). Conservation of an intact vif gene of human immunodeficiency virus type 1 during maternal-fetal transmission. Journal of Virology *72*, 1092-1102.

Yu, Q., Chen, D., Konig, R., Mariani, R., Unutmaz, D., and Landau, N.R. (2004a). APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. The Journal of Biological Chemistry *279*, 53379-53386.

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

Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., and Yu, X.F. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV Vif-Cul5-SCF complex. Science *302*, 1056-1060.

Yu, Y., Xiao, Z., Ehrlich, E.S., Yu, X., and Yu, X.F. (2004c). Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. Genes Dev *18*, 2867-2872.

Zennou, V., and Bieniasz, P.D. (2006). Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates. Virology.

Zennou, V., Perez-Caballero, D., Gottlinger, H., and Bieniasz, P.D. (2004). APOBEC3G incorporation into human immunodeficiency virus type 1 particles. Journal of Virology *78*, 12058-12061.

Zhang, F., Wilson, S.J., Landford, W.C., Virgen, B., Gregory, D., Johnson, M.C., Munch, J., Kirchhoff, F., Bieniasz, P.D., and Hatziioannou, T. (2009). Nef proteins from simian immunodeficiency viruses are tetherin antagonists. Cell Host Microbe *6*, 54-67.

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

Zhang, L., Huang, Y., Yuan, H., Tuttleton, S., and Ho, D.D. (1997). Genetic characterization of vif, vpr, and vpu sequences from long-term survivors of human immunodeficiency virus type 1 infection. Virology *228*, 340-349.

Zhang, W., Chen, G., Niewiadomska, A.M., Xu, R., and Yu, X.F. (2008). Distinct determinants in HIV Vif and human APOBEC3 proteins are required for the suppression of diverse host anti-viral proteins. PLoS ONE *3*, e3963.

Zhang, W., Du, J., Evans, S.L., Yu, Y., and Yu, X.F. (2011). T-cell differentiation factor CBFβ regulates HIV Vif-mediated evasion of host restriction. Nature *481*, 376-379.

Zhen, A., Wang, T., Zhao, K., Xiong, Y., and Yu, X.F. (2010). A single amino acid difference in human APOBEC3H variants determines HIV Vif sensitivity. Journal of Virology *84*, 1902-1911.

Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M.*, et al.* (2002). Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. Nature *416*, 703-709.

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

Zielonka, J., Bravo, I.G., Marino, D., Conrad, E., Perkovic, M., Battenberg, M., Cichutek, K., and Münk, C. (2009). Restriction of equine infectious anemia virus by equine APOBEC3 cytidine deaminases. Journal of Virology *83*, 7547-7559.

Appendix 1: Permissions

**CAMBRIDGE UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS**

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

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

License Number	2873091316007
License date	Mar 20, 2012
Licensed content publisher	Cambridge University Press
Licensed content publication	Expert Reviews in Molecular Medicine
Licensed content title	Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics
Licensed content author	John S. Albin and Reuben S. Harris
Licensed content date	Jan 22, 2010
Volume number	12
Issue number	-1
Start page	0
End page	0
Type of Use	Dissertation/Thesis
Requestor type	Author
Portion	Full article
Author of this Cambridge University Press article	Yes
Author / editor of the new work	Yes
Order reference number	
Territory for reuse	World
Title of your thesis / dissertation	Human Immunodeficiency Virus Evasion of APOBEC3 Restriction Factors
Expected completion date	Jun 2012
Estimated size(pages)	300
Billing Type	Invoice
Billing address	333 8th Street SE

Print This Page

AMERICAN SOCIETY FOR MICROBIOLOGY LICENSE
TERMS AND CONDITIONS

Mar 20, 2012

This is a License Agreement between John S Albin ("You") and American Society for Microbiology ("American Society for Microbiology") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by American Society for Microbiology, and the payment terms and conditions.

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

License Number	2873091061552
License date	Mar 20, 2012
Licensed content publisher	American Society for Microbiology
Licensed content publication	Journal of Virology
Licensed content title	Long-Term Restriction by APOBEC3F Selects Human Immunodeficiency Virus Type 1 Variants with Restored Vif Function
Licensed content author	John S. Albin, Guylaine Haché, Judd F. Hultquist, William L. Brown, Reuben S. Harris
Licensed content date	Oct 1, 2010
Volume	84
Issue	19
Start page	10209
End page	10219
Type of Use	Dissertation/Thesis
Format	Print and electronic
Portion	Full article
Order reference number	
Title of your thesis / dissertation	Human Immunodeficiency Virus Evasion of APOBEC3 Restriction Factors
Expected completion date	Jun 2012
Estimated size(pages)	300
Billing Type	Invoice
Billing address	333 8th Street SE
	#111
	Minneapolis, MN 55414
	United States
Customer reference info	

347

American Society for Biochemistry and Molecular Biology

To whom it may concern,

It is the policy of the American Society for Biochemistry and Molecular Biology to allow reuse of any material published in its journals (the Journal of Biological Chemistry, Molecular & Cellular Proteomics and the Journal of Lipid Research) in a thesis or dissertation at no cost and with no explicit permission needed. Please see our copyright permissions page on the journal site for more information.

Best wishes,

Sarah Crespi

American Society for Biochemistry and Molecular Biology
11200 Rockville Pike, Rockville, MD
Suite 302
240-283-6616
JBC | MCP | JLR

From: "Werf van der, Nel, Springer SBM NL" <Nel.vanderWerf@springer.com>
Subject: **RE: Copyright Permission for John Albin**
Date: April 10, 2012 4:29:43 AM CDT
To: "Albi0028@umn.edu" <Albi0028@umn.edu>
Cc: "Smilios, Arthur, Springer US" <Arthur.Smilios@springer.com>
▶ 1 Attachment, 3.4 KB

Dear Mr. Albin,

Thank you for your email.

You may post your thesis in the University's Digital Conservancy, however, our material may not be separately downloadable from the thesis. Please use the one year embargo.

Best regards,

Nel van der Werf (Ms)
Springer
Rights and Permissions

(Please note that I do not work on Fridays)

Van Godewijckstraat 30 | P.O. Box 17
3300 AA Dordrecht | The Netherlands
tel +31 (0) 78 6576 298
fax +31 (0)78 65 76-377
Nel.vanderwerf @springer.com
www.springer.com

From: Smilios, Arthur, Springer US
Sent: Monday, April 09, 2012 07:58 PM
To: Werf van der, Nel, Springer SBM NL
Subject: Copyright Permission for John Albin

Hello Nel,

Thank you for your kind assistance, regarding John Albin and his request for permission. He has furthered his request, which I am pasting below. Again, I appreciate your assistance. Do not hesitate to contact me if you have any questions.

Sincerely,

Arthur.

Hello,

Thank you for your response. I would like to reapply for permission, however, for the following reason. Digital submission of my thesis to the University of Minnesota Graduate School requires me to participate in the University's Digital Conservancy, which makes dissertations freely available to the public. May I have permission to submit the portions of my thesis derived from this book chapter to the University of Minnesota Digital Conservancy?

I am planning to embargo my thesis for at least one year, and I can extend that to two, which would prevent the copyrighted sections in question from becoming available before the Springer volume. Moreover, I would point out that I do not intend to use the chapter verbatim; it's simply more efficient for me to use sections I've already written where appropriate in combination with other materials for which I have permission as opposed to rewriting everything.