BIOCHEMICAL AND BIOPHYSICAL INVESTIGATIONS OF HIV-1 VIRAL INFECTIVITY FACTOR AND ITS INTERACTION WITH HUMAN APOBEC3G

by

DAYONG ZHOU

(Under the Direction of John Patrick Rose)

ABSTRACT

The human immunodeficiency virus type-1 (HIV-1) Virus Infectivity Factor (Vif) mediates the degradation of a cellular antiviral factor, Apolipoprotein B mRNA-editing enzyme-catalytic, polypeptide-like 3G (APOBEC3G), by serving as an adaptor that bridges the APOBEC3G and the E3 ubiquitin ligase complex. Human APOBEC3G is a potent inhibitor of retroviruses and functions to deaminate cytidines to uridines in HIV-1 ssDNA during viral replication. Thus, degradation of APOBEC3G neutralized one of the body's main defenses against HIV. The characterization of the Vif - APOBEC3G interaction is important to understand how Vif functions in suppressing APOBEC3G activity, which could lead to new and better HIV drugs and therapies. However, the study of the Vif - APOBEC3G complex has been severely hindered by the fact that both Vif and APOBEC3G proteins have proved to be very difficult to express and purify in large amounts.

The goal of the work presented in this dissertation is twofold: (1) to address the current bottleneck for Vif and APOBEC3G protein production by developing expression and purification protocols that are capable of producing milligram quantities of both proteins and (2)

characterize the Vif protein and the Vif-APOBEC3G complex using biochemical and biophysical methods.

Milligram quantities of Vif were produced using the bacterial expression system. The protein has been characterized using a number of biochemical/biophysical techniques and suggest that Vif is an unstructured protein under native conditions. Circular dichroism (CD) indicated a random coil structure with few secondary structural elements. Nuclear magnetic resonance (NMR) produced a spectrum characteristic of an unstructured protein. Small-angle X-ray scattering (SAXS) studies suggested that Vif forms multimers with a compact core in solution. Overall, these results support the idea that Vif is an intrinsically unstructured protein in solution.

The full-length APOBEC3G was produced using a baculovirus expression system. Both monomeric and dimeric APOBEC3G forms were observed providing the first direct evidence for APOBEC3G dimerization. The binding kinetics of Vif-APOBEC3G interaction was determined by surface plasmon resonance (Biacore). Interestingly, full-length Vif was observed to interact with APOBEC3G dimers (K_D=0.2 nM) with an affinity that is 1000-fold greater than that observed for the APOBEC3G monomers (K_D=200 nM).

INDEX WORDS: Biophysics, HIV-1, Intrinsically Unstructured Protein, Vif, NMR, CD, SAXS, APOBEC3G, Surface Plasmon Resonance, AlphaScreen, Multimerization, Dimerization

BIOCHEMICAL AND BIOPHYSICAL INVESTIGATIONS OF HIV-1 VIRAL INFECTIVITY FACTOR AND ITS INTERACTION WITH HUMAN APOBEC3G

by

Dayong Zhou

B.S., Nanjing Agricultural University, China, 1998

M.S., Shanghai Ocean University, China, 2003

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2011

BIOCHEMICAL AND BIOPHYSICAL INVESTIGATIONS OF HIV-1 VIRAL INFECTIVITY FACTOR AND ITS INTERACTION WITH HUMAN APOBEC3G

by

DAYONG ZHOU

Major Professor:	John Patrick Rose
Committee:	Bi-Cheng Wang
	Jeffrey Urbauer
	Robert Jeffrey Hogan

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2011

DEDICATION

I dedicate this work to my parents and brother for their love and encouragement.

ACKNOWLEDGEMENTS

I would like to gratefully thank Dr. John Rose for his guidance and support during my graduate studies. As my major profosser, his mentorship and encouragement were always invaluable. Also, I owe my deepest gratitude to Dr. Bi-Cheng Wang for giving me the opportunity to pursue my graduate degree. This dissertation would not have been possible without his endless support. I would also like to give my many thanks to all my committee members: Dr. Jeffrey Urbauer and Dr. Robert Jeffrey Hogan for their support and advice in my research.

I am grateful for the support I received from Dr. Zhijie Liu, for his guidance and assistance when I was working in his lab. I would also like to acknowledge Dr. Hao Xu for his helpful suggestion and advice throughtout my research. I would also like to thank Dr. Gary Newton for his help during my preparation of this dissertation. Again, this dissertation would not have been possible without the aid of Lily Li. I would like to deeply thank her for her selfless help during my graduate studies.

Finally, I would like to thank all the members in my lab and the Department of Biochemistry & Molecular Biology for their encouragement and great help in the research.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

In 1983, the isolation of the human immunodeficiency virus (HIV) [1] represented the beginning of almost thirty years of intense worldwide battle against the virus. HIV causes Acquired Immune Deficiency Syndrome (AIDS), the end-stage disease of HIV infection, that was first identified as a pathologically distinct entity in June 1981 [2]. Even prior to the discovery of HIV, various studies had provided the first insights into the pathogenesis of HIV infection: induced immune dysfunction involving CD4+ T cells, macrophage depletion and systemic immune dysregulation and failure [3-5]. Our understanding of the pathogenesis of HIV/AIDS has evolved dramatically over recent years primarily due to the rapid development of structural and molecular biology. Progress in defining viral targets and understanding in detailed mechanism of viral infection and replication at the molecular and cellular levels have lead important discoveries in HIV pathogenesis and have contributed to the development of effective treatments for AIDS.

Transmission of HIV is achieved by sexual activities, by transfusion with contaminated blood products and by intravenous drug abuse when contaminated needles are shared [6]. Because of the difficulty of early identification of infected individuals and the absence of symptoms during the disease progression, HIV has been found in every country and has moved from people engaging in risky behaviors to more general populations [7]. Presently, HIV-1

pandemic still shows no sign of slowing down. WHO/UNAIDS estimates that 33.3 million people were living with HIV-1 infection at the end of 2009, with 2.6 million people becoming newly infected. The number of annual HIV/AIDS-related deaths worldwide is estimated to be 1.8 million in 2009 alone [8]. In spite of the increasingly heroic efforts in basic research and global collaborations, development of an effective vaccine and/or cure is still a long way from success.

THE STRUCTURE AND GENOME OF HIV-1

The structure of HIV-1, as revealed by electron microscopy, shows that the virus has a cone-shaped viral capsid located in the center of the virus, composed of approximately 2000 copies of the viral p24 Gag protein (Figure 1.1). Inside this capsid are two identical copies of single-strand HIV-1 RNA genome stabilized as a ribonucleoprotein complex with about 2000 copies of the nucleocapsid (NC) proteins. Each copy of the RNA is also tightly bound to the enzymes (HIV protease (PR), HIV reverse transcriptase (RT) and HIV integrase (IN)) that are closely associated with viral development and infection [9]. The association of the nucleocapsid proteins with the HIV genomic RNA protects the genome from nuclease digestion, whereas the reverse transcriptase and integrase are responsible for viral infection. Surrounding the inner surface of the membrane of the virus is approximately 2000 copies of the myristoylated HIV-1 matrix protein (MA) p17. The MA p17 protein is proteolytically excised from the N-terminus of the HIV Gag polyprotein [10] and forms the HIV-1 viral matrix. The viral matrix is vital for the structural stability and integrity of the virion. Like other retroviruses, HIV-1 is enveloped by a lipid bilayer that is derived from the membrane of the host cell. Anchored on the viral membrane are HIV envelope glycoprotein (Env) spikes composed of surface envelope glycoprotein (SU;

gp120) and transmembrane envelope protein (TM; gp41). A three-dimensional model of HIV-1 spike structure has been generated using cryo-electron microscopy [11]. In this model, the viral envelope spike is composed of three copies of gp120-gp41 heterodimers forming a heterodimer trimer. The gp120 protein forms the 'head' of each gp41 subunit of the trimeric Env spike, whereas the base of the spike contains three gp41 subunits arranged in a tripod-like conformation that project from the trimer head [12]. The structure of the viral envelope spikes is of particular importance because they initiate the virus-cell attachment and infection.

At the onset of infection, HIV-1 binds specifically to the cells expressing the surface receptor CD4, which is a protein with a broad range of immune recognition and signaling functions. The binding of viral pg120 with N-terminal immunoglobulin domain of CD4 is sufficient for virus-cell attachment but not for infection. Unlike other retroviruses, HIV-1 requires additional cell surface receptors to promote infection and fusion of the virus with host plasma membrane. One of several chemokine receptors on the surface of the host cell membrane, such as CXCR4 and CCR5 serve as co-receptors in conjunction with CD4 to trigger the initiation of HIV-1 membrane fusion and subsequent infection [14, 15]. In this process, once the HIV-1 gp120 envelope protein interacts with CD4, gp120 undergoes a conformation change, exposing its chemokine receptor binding domains and allowing them to interact with either CXCR4 or CCR5 co-receptor [16]. The binding of a co-receptor allows the virus to form a more stable and specific virus-cell attachment. After HIV-1 has fused with the target cell membrane, the HIV-1 genomic RNA and associated enzymes, including reverse transcriptase, integrase, and protease, in the form of a ribonucleocapsid complex, are injected into host cytoplasm. The viral RNA undergoes reverse transcription shortly after using its own associated reverse transcriptase, and generates the double-stranded viral DNA. The viral DNA is then transported into the nucleus,

where it is integrated into the host cell genome with the help of associated viral integrase, thereby giving rise to the HIV-1 provirus. The expression of the HIV-1 provirus is initiated by the binding of certain host transcription factors, such as NF-κB, NFAT to the viral long terminal repeat (LTR) [17]. The production of viral mRNAs is sequential, the first viral mRNAs produced are coding for the Tat (Trans-Activator of Transcription), Rev (Regulator of Virion Expression), and Nef (Negative Factor) accessory proteins [18-20]. Subsequently produced are viral structural proteins that allow the assembly of new HIV-1 virions. The newly assembled virions are then released by budding from the host cells and can subsequently reinitiate the retroviral life cycle by infecting other target cells (Figure 1.2).

One remarkable feature that distinguishes the HIV-1 from other retroviruses is the complexity of its genome (Figure 1.3). Most retroviruses contain only three genes – *gag*, *env* and *pol* [22]. The *gag* and *env* genes code for the viral nucleocapsid and surface envelope proteins, respectively, whereas the *pol* gene encode the reverse transcriptase and other enzymes that are responsible for retroviral replication. However, the size of HIV-1 RNA genome is 9.8kb that not only contains three common structural genes – *gag, env,* and *pol*, but also six additional accessory genes (*vif, vpu, vpr, tat, rev,* and *nef*). It is the individual but concerted functions of these accessory gene products that leads to its notorious pathogenesis of HIV infection (Table 1.1). However, from a therapeutic standpoint, the complexity of HIV-1 genome might also be the Achilles heel of the virus [23]. Tremendous efforts have been exerted to search for the antagonists specific for these viral gene products since the discovery of HIV-1. Although numerous breakthroughs have been achieved, the understanding of the expression kinetics of HIV-1 genes is of particular importance in designing antiviral drugs, developing effective vaccines, and treating AIDS progression.

HIV-1 REPLICATION

We now know that the expression of HIV-1 genes is an intricately controlled interplay that involves the chromosome-integrated provirus, host cellular transcription factors, and viral encoded activators. It is now clear that early HIV-1 gene expression is initiated by binding of the host transcription factors, such as NF-κB, NFAT, to the viral LTR to induce a low, basal level of HIV-1 gene expression. The initial viral mRNAs that enter the infected cell cytoplasm are exclusively in the form of the small multiply spliced transcripts, around 2-kb (kilobase) in size [25]. These mRNAs exclusively encode the viral regulatory proteins Tat, Rev, and Nef. The early expression of these accessory proteins has a great impact on the regulation of further HIV-1 replication. The viral Tat protein, once a threshold amount is expressed, can potently increase both the synthesis of HIV-1 RNA and the expression of all HIV-1 genes by establishing a powerful positive-feedback loop [26-28]. Much effort has been focused on the understanding of functions of Tat and the mechanism by which it trans-activates viral gene expression. Depending on the HIV subtype, Tat is between 86 and 101 amino acids containing several distinct domains: a N-terminal activation domain, a Cys-rich region, a highly basic RNA binding domain, and a nuclear localization signal sequence. Tat acts on a RNA stem-bulge-loop structure, known as transactivation response region (TAR) that is present at the 5' end of all viral mRNAs. Tat and other cellular factors, such as Cyclin T1 [29], bind to the TAR, where Tat recruits the human positive-transcriptional elongation factor b (P-TEFb) complex to TAR, resulting in a dramatic upregulation of viral gene transcription [30, 31].

Nef, the second regulatory protein expressed early in the course of HIV-1 infection, is a 205 amino acid residue, myristoylated phosphoprotein. It plays a crucial role in HIV-1 replication and pathogenesis [32, 33], since loss of Nef function gives rise to a decreased viral

replication, and a delayed or abolished progression towards AIDS [34, 35]. The major function of Nef is to enhance virus infectivity by the down regulation of cell surface CD4 molecules and MHC-I (major histocompatibility complex-I) molecules [36-38]. Although the expression of co-receptor CD4 molecules on the surface of cells is essential to HIV-1 infection continued expression of CD4 post-infection poses several problems in the viral replication cycle. Since newly synthesized CD4 molecules are capable of forming complexes with the Env precursor gp160 in the endoplasmic reticulum (ER) they can prevent gp160 processing into gp120 and gp41 or block the transport of gp120 and gp41 to the site of virion assembly [39]. In addition, surface CD4 expression can lead to the internalization of newly synthesized virions at the plasma membrane and/or superinfection of cells by cell-free or cell-associated virus [40]. However, HIV-1 has evolved a distinct mechanism, mediated by the viral Nef and Vpu proteins (Vpu will be discussed later), to down-regulate CD4 expression in the host cells after infection. The other major target of Nef is the MHC-I molecule, which can present small antigenic peptides from HIV-1 (and all other) antigens to cytotoxic T lymphocytes leading to the destruction of the infected cells displaying foreign antigens. This mechanism is the basis of acquired anti-viral immune response [41]. By disruption of MHC-I, Nef efficiently protects HIV-1 infected cells from being killed by cytotoxic T lymphocytes, and ensuring the survival of infected cells *in vivo* [42, 43]. Interestingly, Nef down-regulates these two surface receptor proteins by different intracellular trafficking mechanisms. To disrupt CD4 that has been transported to the cell surface, Nef interacts with the cytoplasmic tails of CD4 and mediates the endocytosis of CD4 molecules through clathrin-coated pits [44, 45]. Nef-mediated prevention of export of new MHC-I results from the disruption of normal transport of MHC-I from the Trans-Golgi Network

(TGN) to the cell surface, inducing the internalization and accumulation of MHC-I molecules in endosomal vesicles, and their subsequent degradation [46].

The process of HIV-1 transcription is intrinsically dynamic, producing a large amount of viral RNAs. Besides the early transcribed small multiply spliced mRNAs, which encode Rev, Tat, and Nef, there are two other major classes of viral RNAs left for transcription and translation. These are 1) full-length unspliced RNAs that function as the mRNAs for the Gag and Pol precursors, and are packaged into progeny virions as genomic RNA, and 2) partially spliced mRNAs, around 5-kb in size that encode the Env surface proteins, and Vif, Vpu, Vpr accessory proteins. However, the accumulation of these RNAs in the host nucleus poses a problem for HIV-1, because they need to be transported to cytoplasm for translation into functional proteins. This problem has been overcome through the interaction with a novel viral protein – Rev (Regulator of expression of viral proteins). Rev, a 19kDa (116 amino acid) phosphoprotein, functions at the viral cis-acting RNA regulatory element (RRE) and provides time-dependent regulation during viral replication [47, 48]. RRE is a highly structured RNA element (~250 nucleotides) that is present in all unspliced and partially spliced HIV-1 RNAs. There are two functional domains in Rev protein: an Arg-rich sequence that is required for viral RNA binding and nuclear localization, and a hydrophobic, Leu-rich motif that mediates nuclear export [48]. The Rev monomer initially binds to RRE and then, through cooperative protein-protein and protein-RNA interactions, forms multimers with other Rev molecules. These Rev multimers eventually bind the RRE at a ratio of up to eight Rev molecules per RRE [49]. The Rev-RRE complexes are capable of interacting with the cellular CRM1-dependent nuclear export machinery. As a result, the RRE-containing RNAs are transported to the cytoplasm under the

direction of Rev, then Rev shuttles back to the nucleus for recycling after unloading viral RNAs, using an intrinsic nuclear localization signal [13].

Late Expression of HIV-1 Structural and Enzymatic Proteins

As mentioned above, the early stage of HIV-1 gene expression is characterized by expression of viral regulatory proteins – Tat, Nef, and Rev. These proteins perform their functions with a surprisingly high efficiency on viral transcription activation, transportation, and protection, ensuring the subsequent effective production of other viral structural and enzymatic proteins for the assembly of infectious HIV-1 virions. After early expression of these three accessory proteins, several major HIV-1 structural and enzymatic proteins need to be synthesized from their precursors, including: (1) the Gag polyprotein precursor, which consists of matrix (MA), capsid (CA), nucleocapsid (NC) and p6, along with two spacer peptides SP1 and SP2, (2) the Gag-Pol polyprotein precursor, which is generated by a -1 ribosomal frameshift at 3' end of Gag and encodes the viral enzymes including PR, RT and IN, (3) the Env glycoprotein precursor, gp160, which is cleaved into the gp120 and gp41 by cellular protease in the Golgi apparatus [24, 50], and (4) several other accessory proteins: including Vif (Viral Infectivity Factor), Vpu (Viral Protein U), and Vpr (Viral Protein R).

HIV-1 Assembly

Following the expression of the full complement of viral proteins, the HIV-1 assembly process takes place. Initially, HIV-1 assembles and buds from infected cells as an immature, noninfectious virion characterized by a layer of self-associated Gag proteins that are bound with the inner viral membrane. The immature virion particles then undergo a dramatic rearrangement

leading to change in structure to become morphologically distinct mature, infectious virions upon budding [21]. The major machinery in the HIV-1 assembly is Gag precursor polyprotein – $Pr55^{Gag}$, from which all the structural proteins of HIV-1 are all derived. The Gag precursor proteins are expressed with an extended conformation consisting of the MA, CA, NC and p6 protein domains [51]. Each of these proteins has a distinct function,during the viral assembly process. Interestingly, the question of where HIV-1 assembly takes place in the cell has been a subject of confusion and controversy for many years. It is now clear however, that the major site of HIV-1 assembly is the plasma membrane (PM) [52]. MA domain at the Gag N-termius directly binds to the inner surface of plasma membrane, where a myristoyl modification located within a highly conserved patch of basic residues promotes the tight membrane binding of MA by inserting the myristoyl group into the lipid bilayer [53]. The membrane binding of MA domain actively permits the Gag-Gag multimerization to take place allowing the further assembly of progeny virions [54]. Several studies have indicated that the C-terminus of the Gag CA domain plays an essential role in Gag-Gag multimerization as well [54, 55]. This C-terminal dimerization domain of CA contains the "major homology region" which is highly conserved among many retroviruses. Mutations within this domain often result in the defects in virus assembly, suggesting the requirement of this region in Gag-Gag interactions [56].

As noted above, each HIV-1 particle contains two single-stranded copies of genomic RNA. The cis-acting sequence that directs RNA encapsidation, known as the packaging signal is located between HIV-1 RNA 5' LTR and the initiation codon of the *gag* gene [57]. RNAs that lack the packaging signal cannot be efficiently encapsidated into virus particles. The specific encapsidation of HIV RNAs into virus particles is mediated by the binding between the RNA packaging signal and the NC domain of Gag (with a dissociation constant ~100nM) forming the

innermost layer of immature virion; the high-resolution solution structure of HIV-1 NC protein revealed that NC contains two CCHC-type zinc finger domains separated by a basic linker segment and flanked by N- and C-terminal tails [58]. The basic residues confer the non-specific nucleic acid binding ability, whereas the two zinc finger domains enable NC to interact with HIV-1 RNA packaging signal more specifically. It has been proposed that NC recognizes the specific sequence of packaging signal with a higher affinity, followed by a nonspecific, sequence-independent binding of NC to the viral genomic RNA [13, 59].

Env Transport and Incorporation

A key step in the production of mature infectious HIV-1 virions is the incorporation of the Env glycoprotein complex. As mentioned above, the HIV-1 Env glycoproteins are synthesized as the Env precursor protein, gp160 in the ER, and then, transported to the cell surface via the secretory pathway. During its trafficking to through Golgi, gp160 precursor is cleaved by a cellular furin or furin-like protease to generate mature surface glycoproteins, gp120 and gp41 [50]. Following cleavage, gp120 and gp41 still remain associated by noncovalent interactions and form the envelope spike in a trimeric complex of heterodimers [12]. When the Env glycoprotein complexes reach the cell surface, they are either rapidly internalized or incorporated into virus particles. The actual mechanism by which the Env glycoproteins are incorporated into virions is still unclear. Two models have been proposed: (1) the passive incorporation model, in which glycoproteins are readily incorporated into virion simply because abundant gp120/gp41 are expressed and present on the cell surface and (2) the Gag–Env interaction model, in which the binding of the gp41 cytoplasmic tail (CT) to the MA domain of

Gag mediates Env incorporation. Since these two models are not mutually exclusive, it is possible that each contributes to Env incorporation to a certain extent [50].

HIV-1 Budding

The last step in the process of HIV-1 assembly is budding – the release of virus particles from the host cell plasma membrane. Although it was thought intuitively that budding is a spontaneous event, it has become increasingly clear that a sequence referred to as "late" or "L" domain in the p6 domain of Gag specifically promotes viral particle budding [60]. Deletion of p6, or mutations within a highly conserved Pro Thr/Ser-Ala-Pro (P-T/S-A-P) motif located near the N-terminus of p6, has been shown to dramatically impair particle release [61]. Although the mechanism by which L domains directs virus release remains unclear, more and more evidence suggest that these domains exert their functions by interacting with host factors involved in the host ubiquitination pathway [62, 63].

Virus Maturation

Up to this point, the immature, noninfectious virion is synthesized in a roughly spherical shell of radially extended Gag molecules. The MA domains are bound to the viral inner viral membrane, and the Gag C-termini project towards the center of the virion. Studies performed using cryo-EM have visualized Gag monomers in immature virions being aligned like spokes on a wheel [64]. The envelope glycoprotein spikes are present on the surface of the particle and two copies of HIV-1 genomic RNA are observed inside the viral core. To become infectious, immature virions must undergo a series of further events, called maturation. Shortly after virus budding from the plasma membrane, the viral protease (PR) that is generated by autoproteolysis

at late stage of viral assembly cleaves the Gag and Gag-Pol polyprotein precursors to generate the mature Gag and Pol proteins. Gag is cleaved to generate MA, CA, NC and p6 proteins and the SP1 and SP2 spacer peptides (Figure 1.4); MA covers the inner leaflet of the viral membrane and is tightly associated with it. CA forms a shell around the viral core, NC coats the genomic RNA protecting it from damage, and p6 engages in the interactions with host factors for virus budding. The functions of SP1 and SP2 spacer peptides are not currently known; however, they appear to regulate rates of cleavage and virus budding [65, 66]. The Pol portion of Gag-Pol is cleaved to produce PR, RT and IN enzymes that are associated with RNA inside the viral core. The most visible morphologic change in HIV-1 maturation is that virion core is converted from doughnut-like to conical shape. As mentioned above, in order to become infectious, the HIV-1 virion must transition from the immature to mature form with the absolute requirement of viral PR. This PR-dependent virion maturation has been the pharmacological target. The inhibition of PR has applied to the treatment of HIV/AIDS patients using PR inhibitors combined with RT inhibitors known as HAART (highly active antiretroviral therapy) [67]. A long-lasting, significant reduction in virus loads, even to undetectable levels, among AIDS patients has been achieved,

ROLES OF OTHER HIV-1 ACCESSORY PROTEINS

In addition to the proteins encoded by other retroviruses (i.e., the products of the gag, pol, and env genes), and the regulatory/accessory proteins (Rev and Tat), HIV-1 encodes several additional proteins with a variety of functions, but many of them are poorly understood. One of the features of HIV-1 that distinguishes it from other retroviruses is the array of ‘‘accessory’’ proteins it produces. The principal function of these accessory proteins is to assist the virus to

evade from various forms of host antiviral resistance. Broadly speaking, the viral accessory proteins changes the subcellular environment within infected cells to ensure viral replication and transmission [68].

Vpu (Viral Protein U)

Vpu is an integral membrane phosphoprotein with a molecular weight around 9 kDa and is comprised of a short N-terminal region, a transmembrane domain (TMD), and a longer cytoplasmic domain (CD) [69]. Vpu is exclusively encoded by HIV-1 but not expressed in HIV-2 and most of other retroviruses, and is expressed from the same mRNA as Env, which ensures the synchronous expression of both proteins. Vpu carries out two well-established functions in the HIV-1 replication cycle: (1) it disrupts CD4 trafficking and mediates it to the proteasome for degradation, and (2) it enhances virion release. As mentioned above, HIV-1 has developed several mechanisms to down-regulate both intracellular and surface expression of CD4 molecules whose degradation is essential for HIV-1 replication. All these mechanisms are mediated by the involvement of two accessory proteins: Nef and Vpu. Interestingly, these proteins regulate the CD4 expression at different cellular compartments through distinct mechanisms. Unlike Nef which functions mostly on mature CD4 molecules present at the plasma membrane by a clathrin-mediated endocytosis and targets them for destruction in lysosomes, Vpu, on the other hand, interacts with newly synthesized CD4 in the endoplasmic reticulum, facilitating the binding of CD4 with the SCF ubiquitin ligase and leading CD4 into the endoplasmic-reticulum-associated degradation pathway for degradation. Vpu is synthesized in ER in tandem with the synthesis of CD4, α-helices in the Vpu cytoplasmic domain binds to a very specific LSEKKT motif within the cytoplasmic domain of CD4 [70]. The binding of CD4

and Vpu is required for subsequent degradation; however, it is not sufficient. Following binding, Vpu is phosphorylated, and this phosphorylation recruits the cellular F-box protein β-TrCP generating a CD4-Vpu-β-TrCP ternary complex on the ER membrane. The β-TrCP recruits other components of the E3 ubiquitin ligase complex to form an SCF complex (Skp, Cullin, F-box containing complex) that mediates the ubiquitination of the CD4 and leads to the complete degradation in 26s proteasome [43, 69]. It is worth noting that Vpu-mediated ubiquitination leads to the degradation of CD4 but not Vpu itself. How the viral Vpu protein escapes the degradation, while it is recognized as a substrate, remains an open question [71].

Early studies showed that virion release from cells infected with Vpu-deficient HIV-1 is reduced in a cell-type-specific manner; Vpu is required for efficient virion release from human HeLa cells and primary T cells, but unnecessary for human and simian cell lines such as 293T, HT1080, COS, CV-1 and Vero [72, 73]. The observation that heterokaryons of HeLa and COS-7 exhibit the phenotype of HeLa has indicated that an inhibitor for virion release exists in HeLa cells and that Vpu antagonizes the inhibitor [74]. In 2008, Neil and colleagues finally identified CD137 as a tetherin that specifically inhibits virion release that is counteracted by Vpu [75]. So far, the mechanism of how Vpu antagonizes tetherin is still unknown, but it could be possible that Vpu interacts with tetherin through its TM domain, recruits ubiquitin ligase, and then degrades tetherin in a manner similar to how Vpu degrades CD4 [76].

Vpr (viral protein R)

Vpr is a multifunctional viral protein that plays significant roles at multiple stages of the HIV-1 life cycle. The specific activities of Vpr include nuclear import of the HIV-1 pre-integration complex (PIC), and the induction of cell cycle G2 arrest. HIV-1 Vpr is a small, 96-

amino acid (14 kDa) protein that appears at two different stages during the virus life cycle: packaged in the virus particles and consequently present in the cytoplasm of newly infected cells [77]. The solution structure of Vpr consists of three bundled α-helices spanning residues 17–33, 38–50, and 55–77, and flanked by flexible, unstructured N- and C-terminal domains that are negatively and positively charged, respectively [78]. One of the unique Vpr activities is to transport the viral PIC between cytoplasm and nucleus. In order to infect the host cells, HIV-1 needs to transport its genomic RNA in the form of viral PIC from the cytoplasm into the nucleus of the target cell. Vpr is believed to be one of the main regulators of HIV-1 nuclear import; other viral proteins involved in the nuclear transport of PIC include MA and integrase (IN) [79]. Although the exact mechanism of Vpr-mediated nuclear transport is still under investigation, recent studies suggest that it might require a direct and/or indirect interaction with importin-α, importin-β, and other nuclear pore proteins. In summary, at least three different models can potentially explain the mechanism of Vpr-mediated PIC nuclear transport. Interactions of Vpr with different cellular proteins may represent different molecular mechanisms in a cell type-dependent fashion [79].

Another unique function of HIV-1 Vpr is to inhibit T-cell proliferation by blocking them in the G2/M phase of the cell cycle, which is commonly known as the G2 arrest [80, 81]. Although the physiological meaning of Vpr-induced G2 arrest, and why HIV would evolve a cell-cycle arrest phenotype are still unknown, it is believed that G2 arrest suppresses human immune response by blocking T-cell clonal expansion and differentiation, and thus maximizes viral replication [82].

Vif (Viral Infectivity Factor)

HIV-1 Vif is a 23 kDa basic phosphoprotein, expressed in the latter stage of the HIV-1 life cycle and conserved among all of the primate lentiviruses. Current knowledge indicates that Vif protein has no established enzymatic activity. Instead, it appears to serve primarily, if not exclusively, as an adapter molecule to mediate the interaction of other viral or host factors.

Initially, Vif was demonstrated to be required for production of infectious progeny virions in some, but not all cell types. In the absence of Vif, virions were estimated to be approximately 1000 times less infectious than the ones produced from wild-type virus [83]. Therefore, the function of Vif is cell species specific and is tightly linked to the nature of the virus-producing cell. Several T-cell lines (e.g. Jurkat, CEM-SS and SupT1) and non-haemopoietic cell lines (e.g. HeLa, 293T and COS) produce infectious HIV-1 virions in the absence of Vif (Δvif HIV-1) and are termed 'permissive'. Conversely, Δvif HIV-1 virions derived from nonpermissive cells, including primary CD4 T cells (e.g. HUT78, CEM) and macrophages, are non-infectious (Figure 1.5) [84]. In other words, Vif is critical for virus replication in primary CD4 T cells and macrophages which are the natural targets of HIV-1 infection, yet entirely dispensable in other T-cell lines. Despite the early identification of permissive and nonpermissive cell types, which supports or inhibits the spread of Δvif HIV-1, respectively the molecular basis for this interesting "Vif phenotype" remained elusive for many years. The mechanism by which Vif allows wild-type HIV-1 virions to spread readily in nonpermissive cells remained unclear as well.

However, some early studies also found that nonpermissive cells support the normal production of Δvif HIV-1 virions, but these virions were unable to productively infect new target cells [85, 86]. This important finding suggested that Vif either overcame the effects of a negative

factor produced in nonpermissive cells or, alternatively, permissive cells express a host Vif analogy that positively influences virion infectivity. This issue was resolved by a fusion cell experiment in which heterokaryons formed between permissive and nonpermissive cells were infected with Δvif HIV-1, which showed that the progeny virions from these heterokaryons were noninfectious [87]. The result that the nonpermissive cell is dominant in heterokaryons strongly supports the hypothesis that nonpermissive cells express an inhibitory factor whose function is somehow countered by Vif. In addition, the heterokaryon experiments clarify the direction of further explanation of Vif function; but the identity of this cellular Vif-sensitive antiviral factor remained a mystery for more than two decades. Recently, Sheehy *et al*. succeeded in identifying human apolipoprotein B mRNA-editing enzyme, catalytic polypeptide- like 3G (APOBEC3G or A3G), as the host Vif-sensitive antiviral factor by screening two genetically related cell lines – nonpermissive CEM T cell and its permissive clone, CEM-SS using a subtractive hybridization assay [88]. In their experiments, the potential cDNA candidates were first verified by differential expression screening and then validated by examining their antiviral activity in single-round infectivity experiments. They observed that A3G was expressed almost exclusively in nonpermissive cell lines and surprisingly, that although A3G was not expressed in permissive cells, the introduction of A3G into permissive cells was found to be sufficient to transform permissive cells into nonpermissive cells [85].

A3G belongs to a family of cytidine deaminases consisting of 11 members: APOBEC-1 (Apo1), APOBEC-2 (Apo2), activation-induced cytidine deaminase (AID), APOBEC-3A, -3B, -3C, -3DE, -3F, -3G, -3H (A3A–H), and APOBEC-4 (Apo4). These enzymes catalyze hydrolytic deamination at the C4 position of the cytidine (or dC) base, converting cytidine to uridine (or dC to dU) on single-stranded (ss) DNA or RNA. All the enzymes in the APOBEC family contain at

least one highly conserved cytidine deaminase domain (CD) [(Cys/His)-Xaa-Glu-Xaa23–28-Pro-Cys-Xaa2–4-Cys]. Four APOBEC enzymes, A3G, A3F, A3B, and A3DE possess two catalytic domains located at N- and C- terminus of the protein bridged by a linker sequence, these repeated CDs appear to be a tandem duplication of the ancestral gene [89]. The residues in the CD coordinate one Zn ion, which is essential for the cytidine deamination activity. Interestingly, only the C-terminal CDs of A3G and A3F are enzymatically active, whereas both N- and C-terminal CDs of A3B are enzymatically active [90, 91]. The reason for the varied CD activity for this enzyme is so far unknown.

MULTIFACETED ANTIVIRAL ACTIVITY OF A3G

Deaminase-Dependent Antiviral Mechanisms

In nonpermissive cells, A3G is effectively incorporated into nascent HIV-1 virions in the absence of Vif through interaction with the nucleocapsid region of the HIV-1 Gag polyprotein. The encapsidated A3G appears to bind to the viral core and therefore, is effectively introduced into the host cells upon virion fusion.

During HIV-1 reverse transcription, A3G selectively recognizes newly reverse-transcribed single-stranded viral DNA (minus/first strand of viral DNA) as substrates that are subjected to A3G cytidine deamination. A3G then extensively catalyzes dC-to-dU mutations on the minus-strand viral DNA. These mutated viral DNAs contain excessive dUs that are subjected to the degradation by DNA repair enzymes and uracil DNA glycosylase [92]. A small amount of viral minus-strand DNAs seem to survive the attack and continues to serve as the template for the viral plus strand DNA synthesis. The dUs on the minus strands promote result in dA misincorporation, a G-to-A hypermutation that exceed 10% of all dG nucleotides in viral DNA

[85]. The incorrect translation termination codons and the altered viral open reading frames introduced by the hypermutation disrupts viral genetic integrity, which can be viewed as a form of error catastrophe [93] that is excessive and sufficient to inhibit the viral infection and replication (Figure 1.6).

Deaminase-Independent Antiviral Mechanisms

Although A3G can potently extensive mutations to HIV-1 DNA, and the antiviral function of A3G is clearly associated with this cytidine deaminase activity, recent studies have revealed that A3G can also exert deaminase-independent antiviral activity [90, 95]. Mutagenesis of highly conserved amino acid residues at in the active site of the A3G C-terminal CD2 abolishes deaminase activity, but these deaminase-deficient mutant enzymes can still inhibit virus infection and replication to a comparatively high level and reduce the amount of HIV-1 DNA produced in new target cells [90]. Many studies have investigated the mechanism of deaminase-independent antiviral activity of A3G. The experimental results obtained suggest that A3G can interfere with primer tRNA annealing, minus- and plus-strand DNA transfer, primer tRNA progression, provirus formation, and viral DNA elongation and integration. Among thses, inhibition of viral cDNA synthesis by impeding RT elongation and interfering viral DNA integration [94, 96] is the primary cause of A3G deaminase-independent antiviral activity in HIV-1.

Importantly, the dominance of deaminase-dependent or deaminase-independent antiviral activity of A3G may vary in different cell lines under different circumstances. This concept is derived from the observation of a remarkable difference in viral cDNA G-to-A mutation levels from Δvif HIV-1 infected primary blood lymphocytes (PBLs) and human T-cell line H9 (both of

which are refractory to HIV-1) [97, 98]. Thus, A3G probably employs a dual strategy involving non-deaminase and deaminase action to achieve its potent antiviral effect. Needless to say, it is important to precisely elucidate the molecular basis for the differential regulation of A3G antiviral activities in different cell types.

VIF-MEDIATED DEGRADATION OF A3G

As described above, A3G alone can effectively inhibit the replication of ΔVif HIV-1 but not the replication of wide type HIV-1. Obviously the major function of Vif is to counteract A3G. As far as the A3G is concerned, its incorporation into budding virions is central to its antiviral activity. Several studies have reported that Vif can reduce ~99% of the A3G incorporation into nascent virions [94, 99], strongly suggesting that Vif inhibits A3G either by direct interference of A3G packaging or by depleting cellular A3G. Although both mechanisms are present, the direct Vif-induced degradation of cellular A3G has been shown to be dominant over blocking A3G encapsidation [93, 100, 101]. More specifically, Vif induces the polyubiquitylation and subsequent proteasome degradation of A3G, therefore depleting the cellular pool of A3G available for encapsidation into nascent virions [93]. To degrade A3G, Vif hijacks a cellular E3 ubiquitin ligase complex consisting of cullin5 (Cul5), elongin B, elongin C, and a ring finger protein (Rbx) (Figure 1.7). Initially, Vif binds to A3G and E3 ubiquitin ligase complex simultaneously, therefore serving as an adaptor molecule to recruit the ligase complex to its substrate – A3G (and possibly Vif itself); then the E3 ubiquitin ligase complex induces the polyubiquitination of A3G and directs it to the 26S proteasome for degradation [102].

Structural Insights into A3G Deamination

A number of studies using mutagenesis and co-immunoprecipitation have been carried out to determine the protein residues involved in the interactions between Vif and A3G. It has been shown that N-terminal region of Vif binds to the N-terminal region of A3G (amino acids 54–124). The binding of HIV-1 Vif to A3G is also critically dependent on an Asp-Pro-Asp motif at positions 128–130 in A3G, especially Asp-128, which is responsible for the species specificity of A3G [103]. The detailed Vif – A3G binding mechanism is still under investigation due to the lack of structure data on the complex and/or A3G N-terminus.

To date, no high-resolution full-length A3G structure is available, whereas the crystal and solution structure of A3G CD2 have been reported (PDB entry: 3E1U and 2JYW) [104, 105]. Basically, the A3G CD2 has an overall β core consisting of five β-strands, surrounded by six α-helices. The deaminase active site comprises residues on helix 2 and 3 that coordinated by a zinc ion. Two loops (loop1 and 3) in the active center (AC) are important for DNA substrate binding and access to the active site. The A3G CD2 crystal structure reveals a substrate groove between AC loop 1 and 3 that may accommodate single-stranded DNA substrate. The DNA could be positioned deep into the pocket of the active site extending out across loop 7 and helix 6 [104]. The active site of A3G CD2 also contains a Zn ion that is coordinated by three residues (2 cysteines and 1 histidine) along with a water molecule. This water molecule, serving as a hydrogen and hydroxyl group donor can be activated to become a Zn-hydroxide coordinated with A3G. The predicted deamination reaction is driven by a direct nucleophilic attack on C4 of the cytidine pyrimidine ring by the zinc-coordinated A3G. The deamination (and resulting oxidation) at position 4 yields a carbonyl group and results in a change from cytidine to uridine [104, 106].

STRUCTURAL UNDERSTANDING OF VIF

Rapid advancement has been made in the studies of the biological functions of A3G in recent years. The structural and functional studies of the interaction between A3G and Hiv-1 Vif have provided a great deal of information for understanding the mechanisms by which Vif mediates the antiviral activity of A3G. However, there are still many questions that remain unanswered mainly due to the lack of structural data of HIV-1 Vif protein and full-length A3G. Importantly, no high-resolution three-dimensional structures for Vif, full-length A3G or the Vif-A3G complex exist. This lack of data is primarily due to the difficulty of producing high levels of soluble recombinant Vif and A3G proteins using either prokaryotic or eukaryotic expression systems. Also, the purification of full-length Vif in large amounts is extremely difficult. Hence, Vif is the only HIV-1 protein for which neither the crystal nor the solution structure is available. To date, the only available structure for Vif is a15-residue fragment (residues 140–155) bound to the Elongin B/C complex (PDB entry: 3DCG) [107]. Other structural information, derived from modeling and indirect experiments is either inconclusive or too controversial to provide the needed details of Vif function and its interaction with A3G.

Overall, Vif consists of several functional domains (Figure 1.8) including (1) an N-terminal domain associated with viral RNA and A3G binding, (2) a central domain that is involved in binding with A3G, the viral protease and Cul5 which contains a His-Cys-Cys-His-zinc binding motif, and (3) a C-terminal domain (CTD) that mediates the interactions with HIV-1 Gag protein, host cellular ElonginC and Vif-Vif multimerization.

It has been suggested that Vif N-terminal region contains the major binding sites for A3G and another important cellular antiviral factor – A3F [108]. Residues 85 to 99 and ^{40}YRHHY^{44} sequence of Vif are important for binding to A3G, whereas the ^{14}DRMR^{17} residues are essential

for binding to A3F. In addition, conserved tryptophan residues (W5, W21, W38 and W89) located in the Vif N-terminus have been shown to be important for Vif suppression of A3G [109]. Other amino acids in Vif are likely to contribute to A3G binding as well, since I9, K22, E45, and N48 substitutions have been shown to suppress Vif function against A3G [108]. These studies suggest that Vif N-terminal domain contains nonlinear binding sites for A3G, and that Vif - A3G interaction appears to be dependent on electrostatic interactions.

HCCH Zn Binding Domain of VIF

As noted above, Vif hijacks the cellular Cul5-E3 ubiquitin ligase complex to degrade A3G. The first A3G residues identified to interact with Cul5 were C114 and C133, but Elongin B/C was shown not to bind these residues [110]. Later, these two cysteine residues were shown to belong to a completely conserved HCCH motif with the sequence H108-X5-C114-X18-C133-X5-H139. The HCCH sequence and spacing are highly conserved among all primate lentiviral Vif proteins, and each individual residue of this HCCH motif was shown to be necessary for interaction with Cul5 [111]. The sequence analysis reveals that the HCCH motif consists of two conserved His/Cys pairs flanking a predicted α-helix that contains a cluster of hydrophobic conserved residues. This motif was shown to coordinate one Zn ion through the universally conserved residues H108, C114, C133 and H139, and to bind directly to Cul5 [108].

SOCS-BOX Domain of VIF

Mutational analysis of Vif has led to the characterization of several distinct domains in Vif for binding with the E3 ubiquitin ligase complex. One of the domains is a highly conserved motif located in the Vif C-terminus, referred to as the SLQ motif. The 144SLQ(Y/F)LA149

sequence resembles a conserved Elongin B/C binding BC-box motif found in SOCS (suppressors of cytokine signaling) proteins and was found to mediate the binding of Vif to Elongin C but not Elongin B [112]. Finally, a conserved PPLP motif (residues 161 to 164) was shown to function in Vif multimerization [113, 114]. This motif was also shown to affect A3G binding, association with Cul5, and interaction with the Pr55gag precursor. The PPLP motif was also found to be important for interaction with the cellular tyrosine kinase HCK [112].

In summary, the multidomain Vif protein recruits several cellular partners to achieve its goal of A3G degradation by forming a A3G-Cul5-E3 ubiquitin ligase complex that targets A3G for degradation by the ubiquitin-proteasome pathway. Vif associates with the Cul5-ElonginB/C complex by binding directly to Elongin C via its SOCS BC-Box motif, and to Cul5 via a conserved zinc-binding HCCH motif. The Vif-associated Cullin5-ElonginB/C ubiquitin ligase complex ubiquitinates A3G and directs it for degradation in proteasome.

The Vif–A3G Interaction as a Pharmacological Target

Recent advances in our understanding of innate immunity of A3G, and the mechanism by which the HIV-1 Vif protein inhibits A3G have provided us new opportunities for the development of novel antiviral drugs.

Thus far, only a few studies have reported attempts to find inhibitors that block Vif-A3G interaction. It has been reported that the proline-rich peptides containing Vif PPLP motif were able to inhibit Vif multimerization and suppress HIV-1 replication in nonpermissive cells [115]. In another study, a membrane-permeable zinc chelator, N,N,N′,N′-tetrakis-(2-pyridylmethyl) ethylenediamine (TPEN), inhibited Vif-Cullin 5 binding and prevented the formation of the E3 ubiquitin ligase, therefore suppressing A3G degradation [116]. Although short peptide inhibitors

have the obvious advantages: they can provide valuable information about Vif-A3G interactions and their function, their therapeutic use as antiviral drugs is limited by their inability to readily cross membrane barriers, low stability and poor bioavailability [117]. On the other hand, in vitro studies have shown TPEN induces apoptosis in human cells and considering that it functions as a zinc chelator, it is very likely to inhibit important cellular enzymes [118].

Undoubtedly, a potent, stable and bioavailable small-molecule inhibitor of Vif would be an exciting medicinal lead. Recently, an exciting development is the identification of a small molecule inhibitor using a cell-based screen of a library of approximately 30,000 compounds; and identified 66 candidate compounds. Among these candidates, the most promising was RN-18 that enhances Vif degradation in the presence of A3G, increases A3G virion encapsidation, and enhances viral DNA hypermutation [119]. The results obtained to date indicate that inhibition of Vif-A3G interaction by either blocking Vif activity or protection A3G from Vif provides a potential target for development of new antiviral therapeutics. Clearly, structural and other studies that elucidate Vif-A3G interaction will undoubtedly provide novel insights and could facilitate rational drug design.

Figure 1.1: Schematic representation of the HIV-1 virion. Locations of the major viral structural and accessory proteins, the membrane, and the viral genomic RNA are indicated. Figure was adapted from [47].

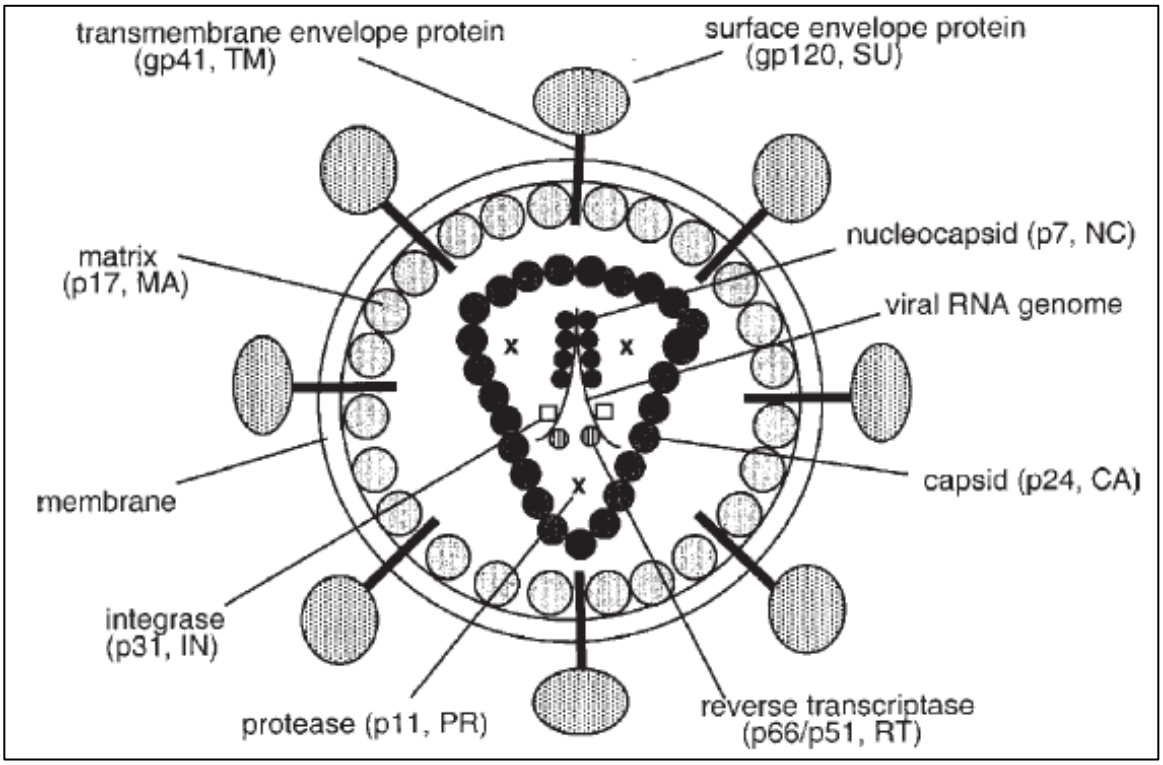


Figure 1.2: The HIV lifecycle. HIV entry involves membrane fusion and releasing of the HIV-1 genomic RNA into the host, where viral RNA is reverse transcribed DNA and integrated into the host genome. The HIV-1 proteins necessary for the production of virions are produced and assembled in cytoplasm. The immature virus particles bud from infected cells followed by maturation becoming mature viruses that continue to infect other cells. Figure was adapted from [50]

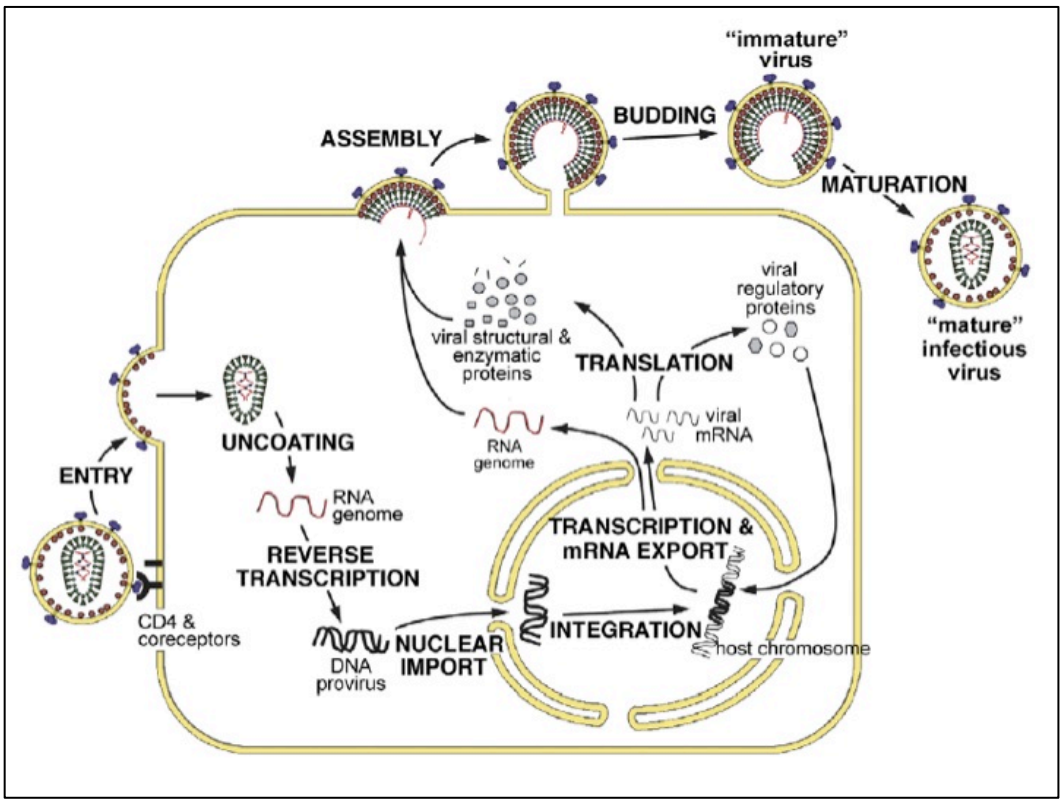


Figure 1.3: Organization of the HIV-1 genome. The gene products encoded by HIV-1 include the Gag proteins matrix (MA), capsid (CA), nucleocapsid (NC) and p6 and spacer peptides SP1 and SP2; the Pol proteins protease (PR), reverse transcriptase (RT) and integrase (IN); the surface Env glycoprotein gp120 and the transmembrane Env glycoprotein gp41; the accessory proteins Tat, Rev, Vif, Vpr, Vpu, and Nef. Also shown are the 5′ and 3′ long terminal repeats (LTRs). Figure was adapted from [48].

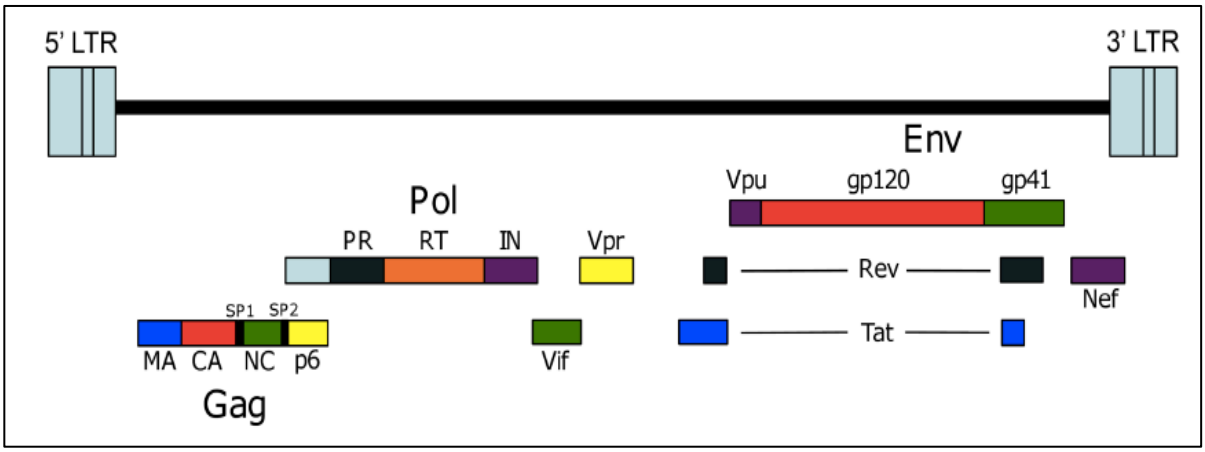


Table 1.1: HIV-1 proteins and their functions

Proteins	Functions
Gag	Building block of virus core Capsid structural protein Matrix protein
Pol	Viral enzymes, including reverse transcriptase, protease, and integrase
Env	gp120 and gp41 envelope proteins
Vif	Disrupts the antiviral activity of human APOBEC3G
Vpu	Helps virion release and promotes CD4 degradation
Vpr	Promotes cell-cycle arrest and regulates nuclear import of the HIV-1 pre-integration complex
Nef	Promotes T cell activation; enhances infectivity; MHC down-regulation
Tat	Enhances viral RNA transcription
Rev	Regulates viral RNA nuclear export

Figure 1.4: Structural model of the extended Gag polypeptide, showing the locations of MA, CA_{NTD}, CA_{CTD}, SP1, NC, SP2, and p6. Unstructured and linker regions are represented by dashed lines. PR cleavage sites are indicated by the arrowheads. Figure was adapted from [50].

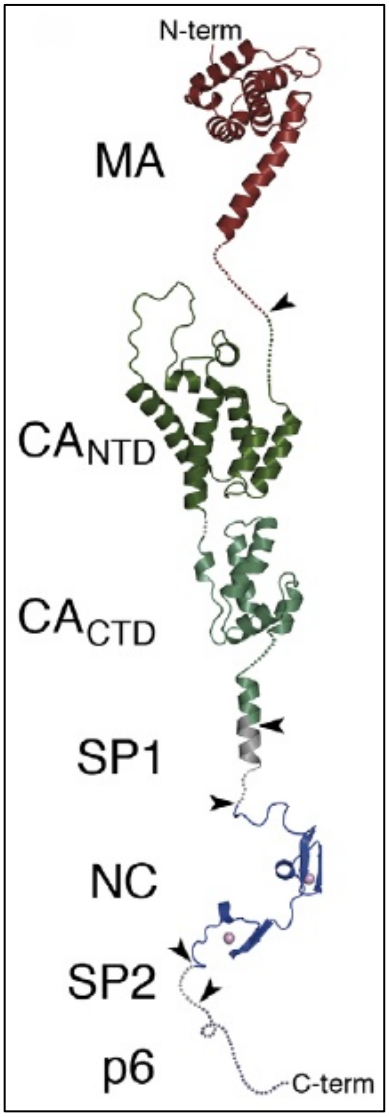


Figure 1.5: The HIV-1 Vif phenotype. ΔVif HIV fail to produce infectious virions in "nonpermissive" cells, which include primary CD4 T cells and macrophages. Conversely, many "permissive" T cell lines and non-hematopoietic cell lines produce infectious virus without the presence of Vif. However, permissive and nonpermissive cells both can produce progeny virions, but virions produced from nonpermissive cells are not able to infect the next target cell. Figure was adapted from [86].

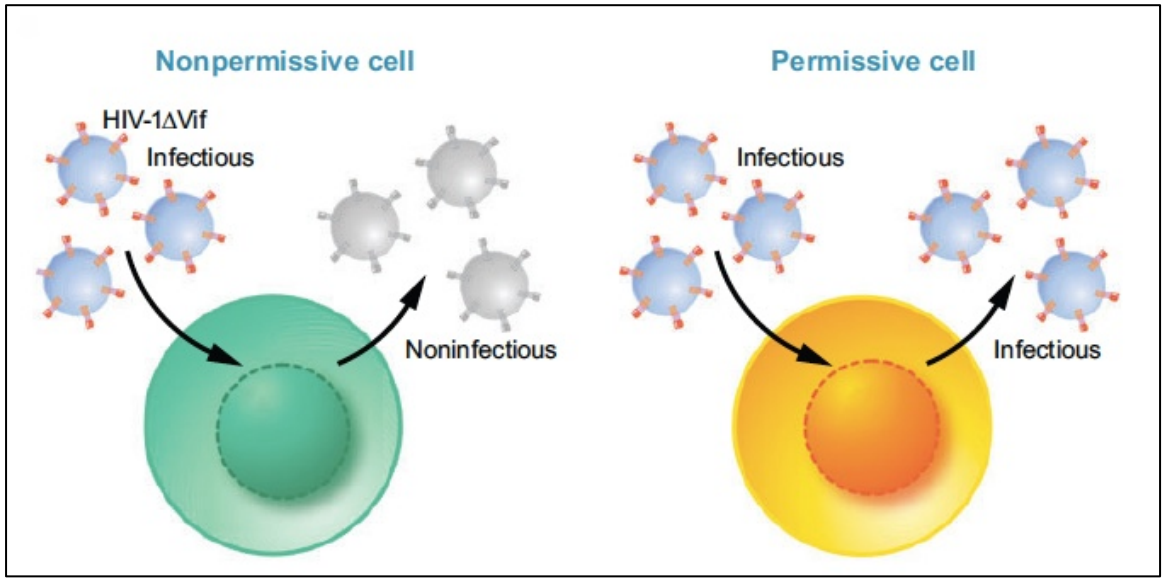


Figure 1.6: Multifaceted antiviral actions of A3G. In the absence of Vif，A3G is not degraded and is effectively incorporated into the budding virus and transferred into the next target cell, where it exerts its deaminase-dependent and deaminase-independent antiviral effects. (1) A3G can inhibit the elongation of the reverse transcriptase in a deaminase-independent manner. A3G binds directly to the viral RNA and therefore impedes the movement of the RT along the RNA template. (2) Most importantly, A3G can trigger massive deamination of dC to dU on the viral minus-strand resulting in G-to-A mutation on the DNA plus-strand, (3) Finally, A3G also inhibits viral DNA integration and provirus formation. Figure was adapted from [95].

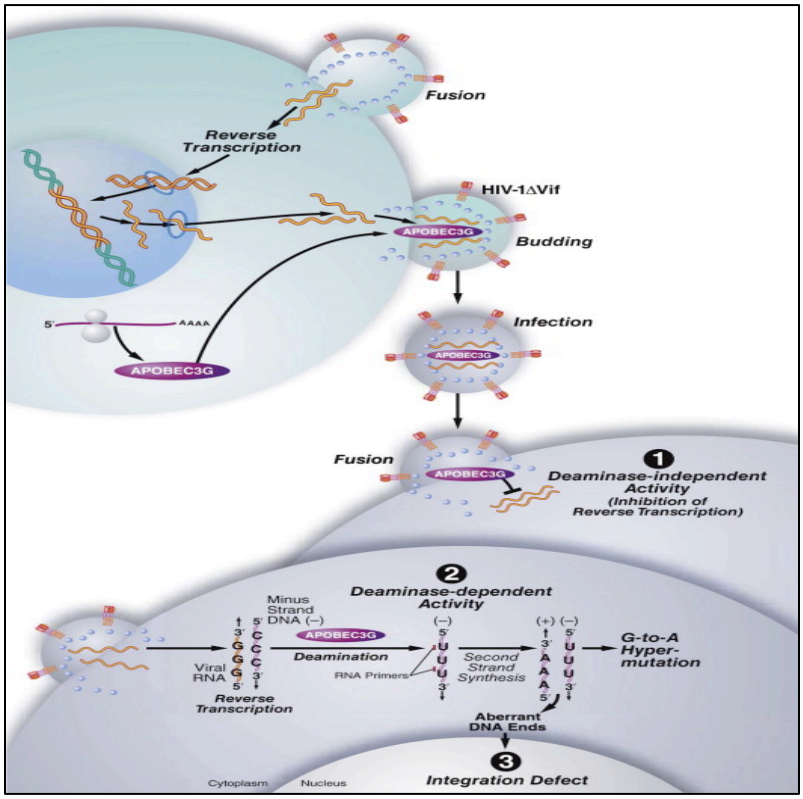


Figure 1.7: Complex of Vif-A3G-E3 ubiquitin ligase. Vif degrades A3G by binding to the A3G and to an E3 ubiquitin ligase complex, simultaneously. This complex consists of cullin5 (Cul5), Elongin B, Elongin C, and a ring finger protein (Rbx). Rbx binds to an unknown E2 ubiquitin-conjugating enzyme (E2). Therefore Vif connects the ligase complex and its A3G protein substrates to initiate polyubiquitination and proteasome-mediated degradation of A3 proteins. Figure was adapted from [95].

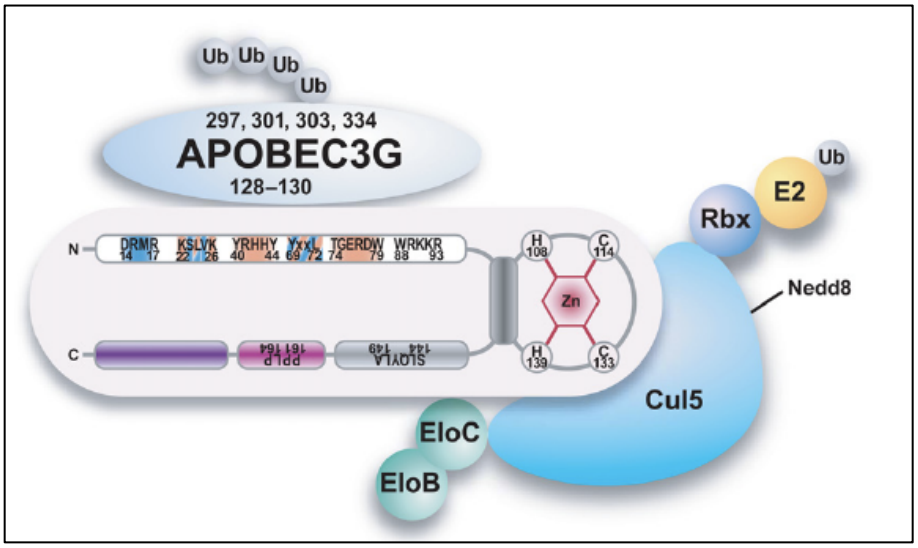
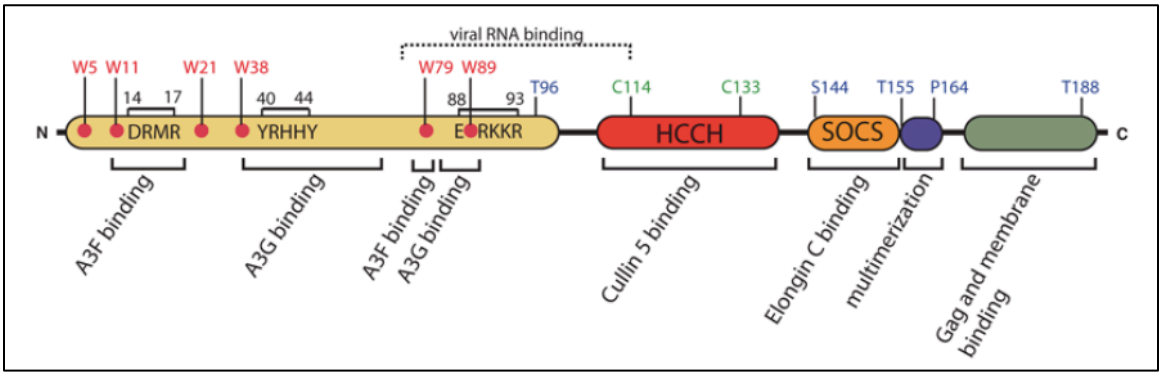


Figure 1.8: The functional domains of HIV-1 Vif. Vif contains several functional domains, including a N-terminal A3G binding domain, a conserved zinc-binding hydrophobic HCCH motif, which is the binding site for Cul5, and a downstream SOCS-box which functions in Elongin C binding. Figure was adapted from [108].



CHAPTER 2

HIV-1 VIF IS AN INTRINSICALLY UNSTRUCTURED PROTEIN

BIOPHYSICAL CHARACTERIZATION OF HIV-1 VIF [1]

[1]Zhou, D., Xu, H., Wang, B.C. and Rose, J.
To be submitted to *The Journal of Biological Chemistry*

ABSTRACT

The human immunodeficiency virus type-1 (HIV-1) Virus Infectivity Factor (Vif) protein mediates the degradation of a cellular antiviral factor, APOBEC3G (A3G), by serving as an adaptor that bridges the A3G and E3 ubiquitin ligase complex. The full-length Vif protein is difficult to purify in large amounts from prokaryotic or eukaryotic expression systems, therefore it is the only HIV-1 protein whose structure remains unknown. Here, we purified Vif protein from inclusion bodies and studied its properties in the unbound state using biophysical techniques and found that Vif is unstructured under physiological conditions. The circular dichroism (CD) spectrum of Vif represented a pattern of random coil with little secondary structure elements. The nuclear magnetic resonance (NMR) spectrum of Vif showed characteristics typical of unstructured proteins. Small-angle X-ray scattering (SAXS) studies suggested that Vif forms multimers with a compact core in solution. Over all, our results obtained from a series of biophysical studies provide direct evidence for the intrinsically unstructured nature of the free full-length HIV-1 Vif in solution.

INTRODUCTION

HIV-1 genome contains nine genes, among which *gag*, *pol* and *env* genes code for structural proteins that are common to other retroviruses. Four Gag proteins, MA, CA, NC, and p6, and the two Env proteins SU gp120 and TM gp41, are building blocks that consist of the viral core and the membrane envelope. The three products of *pol* gene, PR, RT and IN, carry various critical enzymatic functions for effective HIV-1 infection and replication [1, 2]. In addition to the common structural proteins, HIV-1 also expresses six accessory proteins including Rev, Tat, Nef, Vpu, Vpr, and Vif. One of the features that distinguishes HIV-1 from other retroviruses is the array of accessory proteins that assist the virus to evade from various forms of host antiviral resistance. For example, Tat, Nef and Rev proteins perform their functions with a surprisingly high efficiency on viral transcription activation, transportation, and protection ensuring the subsequent effective production of other viral structural and enzymatic proteins for the assembly of infectious HIV-1 virions [3-5]; Vpr is believed to be one of the major regulators of nuclear import of the HIV-1 pre-integration complex (PIC), and induces the cell cycle G2 arrest of host cells [6-8]. Vpu performs two well-established functions including disruption and degradation of CD4, and enhancement of virion release [9, 10]. Broadly speaking, the HIV-1 accessory proteins modify the local environment within the infected cells to ensure viral persistence, replication and transmission [11].

HIV-1 Vif, the subject of this study, is a 23 kDa basic phosphoprotein, expressed in late stage of the HIV-1 life cycle and conserved among all of the primate lentiviruses. Current knowledge indicates that Vif protein has no established enzymatic activities. Instead, it appears to serve primarily, if not exclusively, as an adapter molecule to mediate the interaction of other viral or host factors. Initially, Vif was demonstrated to be required for production of infectious

progeny virions in some, but not all cell types. In the absence of Vif, the virions were estimated to be approximately 1000 times less infectious than the ones produced from the wild-type virus [12]. Therefore, the function of Vif is cell species specific and is tightly linked to the nature of the virus-producing cells. Several T-cell lines (e.g. Jurkat, CEM-SS and SupT1) and non-haemopoietic cell lines (e.g. HeLa, 293T and COS) are able to produce infectious HIV-1 virions in the absence of Vif (Δ*vif* HIV-1) and therefore termed 'permissive'. Conversely, Δ*vif* HIV-1 virions derived from nonpermissive cells, including primary CD4 T cells (e.g. HUT78, CEM) and macrophages, are non-infectious [13]. In other words, Vif is critical for virus replication in primary CD4 T cells and macrophages – the natural targets of HIV-1, yet entirely dispensable in other T-cell lines. Despite the early identification of permissive and nonpermissive cell types (by supporting or not supporting the spread of Δvif HIV-1 respectively), the molecular basis for this implausible "Vif phenotype" remained elusive for many years. The mechanism by which Vif allows wild-type HIV-1 virions to spread readily in the nonpermissive cells remained unclear as well.

In 2002, Sheehy *et al.* succeeded in identifying human apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G (APOBEC3G or A3G), as the Vif-sensitive host antiviral factor [14]. The A3G belongs to a family of cytidine deaminases catalyzing hydrolytic deamination at the C4 position of the cytidine (or dC) base, converting cytidine to uridine (or dC to dU) on single-stranded DNA or RNA. During HIV-1 reverse transcription, A3G selectively recognizes newly reverse-transcribed single-stranded viral DNA (first strand of viral DNA) and extensively catalyzes dC-to-dU deamination. The resulting G-to-A hypermutation is estimated to exceed 10% of all dG nucleotides in HIV-1 DNA [15], which is excessive and sufficient to inhibit the viral infection and replication. A3G alone can effectively inhibit the replication of

ΔVif HIV-1 but not that of wild-type virus as a result of the expression of Vif in wild-type HIV-1. The major function of Vif is to counteract A3G by inducing the polyubiquitylation and subsequent proteasome degradation of A3G. To degrade A3G, Vif hijacks a cellular E3 ubiquitin ligase complex consisting of cullin5 (Cul5), elongin B, elongin C, and a ring finger protein (Rbx). Initially, Vif binds to A3G and E3 ubiquitin ligase complex, thereby serving as an adaptor molecule to recruit the ligase complex to its substrate – A3G. Then the E3 ubiquitin ligase complex induces the polyubiquitination of A3G and directs it to the 26S proteasome for degradation [16]. Rapid advancement has been made in the studies of the biological functions of A3G in recent years. The structural and functional studies of the interaction between A3G and HIV-1 Vif have provided a great amount of information for understanding the mechanisms by which Vif mediates the antiviral activity of A3G. However, there are still many questions that remain unanswered mainly due to the lack of high-resolution three-dimensional structures of HIV-1 Vif. This lack of structural data is primarily due to the difficulty of producing high levels of soluble recombinant Vif and A3G proteins using either prokaryotic or eukaryotic expression systems. Also, the purification of full-length Vif in large amounts is extremely difficult. Hence, Vif is the only HIV-1 protein for which neither the crystal nor the solution structure is available. So far, the only available crystal structure of Vif is the α-helix composed of residues 140–155 within the Elongin B/C complex [17]. Therefore, any structural information either derived from modeling or indirect experiments would be very helpful to further elucidate the function of Vif and its interaction with A3G. Overall, Vif consists of several functional domains as shown in Figure 2.1, including the N-terminal RNA- and A3G-binding domains, the central domain containing a His-Cys-Cys-His- zinc binding motif and Cul5 binding region, and the C-terminal domain (CTD) that mediates the interactions with HIV-1 Gag protein, host cellular Elongin C

and Vif-Vif multimerization. It has been reported that the C-terminal domain residues 141-192, which mediate interactions with both Elongin C and Cul5, are intrinsically disordered [18-20]. The mass spectrometry study on chemical cross-linked Vif also indicated that N-terminal region of Vif is likely to be protected or folded in a compact domain [21].

In this study, we provide direct evidence for the structural features of the HIV-1 Vif. Full-length Vif expressed from *E. coli* was analyzed using biophysical methods including Circular Dichroism (CD), NMR, Thermal shift assay and Small-angle X-ray scattering (SAXS). Vif, in the unbound state, exhibited the characteristics typical of an intrinsically disordered protein under physiological conditions.

EXPERIMENTAL METHODS

Cloning and construction of the recombinant plasmid

HIV-1 Vif was constructed by PCR using pcDNA3.1 vector containing full-length HIV-1 Vif cDNA (a gift from Dr. Tom Hodge at University of Georgia) as the template with 5' primer and 3' primer sequence of GAAAACCTGTACTTCCAAGGCGGGTCAGGTATGGAAAACAGATGGCAGGTGATGA and GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGTGTCCATTCATTGTATGGC, respectively. The PCR product was purified by agarose gel electrophoresis and isolated using a QIAquick Gel Extraction kit (QIAGEN), followed by cloning into pDONR221 plasmid to generate an entry clone. The resulting plasmid, pDONR221-Vif was used as the basis for expression plasmid construction. The *vif* gene was then cloned into Gateway expression vector pDEST527 (N-terminal 6×His tag) by recombination reaction between the entry clone and the

destination vector using LR Clonase (Invitrogen) to generate the expression clone pDEST527-His-Vif. Briefly, pDONR221-Vif and pDEST527 (150 ng each) were added to a 10-μL LR clonase enzyme reaction mixture and incubated for 1 h at 25 °C. *E. coli* TOP10 competent cells were transformed with 1 μL of the recombinant plasmid pDEST527-His-Vif and selected on Luria–Bertani (LB) agar plate containing 100 μg/mL ampicillin. Positive clones were identified by PCR and confirmed by DNA sequencing.

Bacterial expression of Vif fusion protein

For expression of recombinant HIV-1 Vif fusion protein, *E. coli* Rosetta 2 Competent Cells (Novagen) were freshly transformed with expression vector pDEST527-His-Vif (N-terminal 6×His tag). Cells from a single colony were first grown in a small scale of 50 mL LB broth supplemented with 100 μg/mL ampicillin. Cells were cultured at 37 °C with shaking at 250 rpm. Next morning, 20 mL of the overnight culture were used to inoculate a 1 L culture. The expression of recombinant Vif protein was induced by the addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside) to 1 mM when the cell density OD_{600} reached 0.5. After growth for an additional 4 h at 37 °C, cells were harvested by centrifugation for 10 min at 5000*g* at 4 °C and cell pellets were frozen at −80 °C.

For the expression of uniformly ^{15}N- isotope labeled Vif, an aliquot (100 μL) of *E. coli* Rosetta 2 containing pDEST527-His-Vif from glycerol stock was used to inoculate 50 mL LB broth supplemented with 100 μg/mL ampicillin grown at 37 °C overnight. To change the a LB medium to the M9 minimal medium, the cells were pelleted at 5000 g for 10 min, then washed by using 20 mL of M9 medium and pelleted again. The cell pellet was resuspended in a 1000 mL

M9 media containing 2 g/L glucose, 2 mM $MgSO_4$, 0.05 mM $CaCl_2$, 0.01mM $FeCl_3$, 10 mg/L thiamine, 10 mg/L biotin, 100 μg/L ampicillin, and 1 g/L $^{15}NH_4Cl$ as the sole nitrogen source. When the cell density OD_{600} reached 0.5, protein expression was induced by the addition of IPTG to the final concentration of 1 mM. The cells were allowed to grow for overnight at 37 °C. After that, cells were harvested by centrifugation for 10 min at 5000*g* at 4 °C and cell pellets were frozen at −80 °C.

The solubility of recombinant vif was analyzed using CelLytic B Plus Kit (Sigma-Aldrich) following manufacturer's instruction. Briefly, 1 mL cells collected from the induced culture were centrifuged at 13,000 ×*g* for 1 min at 4 °C and the pellets were subsequently resuspended in 50 μL of the CelLytic Working reagent (containing 1 mg/mL hen egg white lysozyme, Protease Inhibitors, and Benzonase) and incubated on ice for 10 min. The insoluble fraction was collected by centrifugation at 13,000 rpm for 10 min at 4 °C, separated from the soluble fraction, and resuspended in 50 μl of SDS-PAGE sample buffer. To evaluate protein solubility, both soluble and insoluble fractions were analyzed by 4–20% SDS–PAGE, followed by Coomassie blue staining.

Purification and refolding of His-Vif fusion protein

For isolation of inclusion bodies, 10g of frozen cell pellets were resuspended in 100 mL of ice cold PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4 \cdot 2H_2O$, and 1.76 mM KH_2PO_4), pH 7.4 containing 5 mM of DTT (Dithiothreitol), 1 mM PMSF (phenylmethylsulfonyl fluoride), Complete Protease Inhibitor Cocktail (Roche), and 1 mg/mL lysozyme. The lysis suspension was incubated for 30 min at room temperature with gentle agitation, followed by

sonication on ice for 3 min. Lysate was cleared by centrifugation at 21,000 ×*g* (JA-25.50 Rotor in Beckman Coulter Avanti® J-E centrifuge) for 30 min at 4°C. The resulting pellet was resuspended in 50 mL wash buffer (containing 50 mM Tris–HCl, pH 8.0, 2% Triton X-100 (v/v), 500 mM NaCl, 1 mM 2-mercaptoethanol, 1 mM PMSF, and 2 M urea) and then centrifuged at 21,000 ×*g* for 20 min at 4°C. This step was repeated three times with 50 mL of the wash buffer without Triton X-100 for the final wash. Following the final centrifugation, the washed inclusion bodies were solubilized in 10 mL of solubilization buffer containing 6 M Guanidine–HCl, 50 mM Tris–HCl, and 1 mM 2-mercaptoethanol, pH 8.0 by incubation at room temperature for 6 h. The solution was cleared from insoluble material by centrifugation at 21,000 ×*g* for 30 min at 4 °C, followed by filtration through 0.45 μm syringe filter.

His-tagged Vif solubilized in 6 M guanidinium chloride (GnCl) was further purified on a 5 mL HisTrap™ HP Immobilized Metal Affinity Chromatography (IMAC) column connected in AKTA Prime system (Amersham–Pharmacia Biotech). The sample was loaded onto the column pre-equilibrated with chromatography buffer A containing 6 M Guanidine–HCl, 50 mM Tris–HCl, 1 mM 2-mercaptoethanol, and 20 mM imidazole pH 8.0. After sample loading, the column was washed with 10 column volumes of buffer A. The His-tagged Vif fusion protein was eluted in a 6-column volume linear gradient of 20–500 mM imidazole in buffer A. Pooled fractions containing His-Vif were either stored at −80 °C or used directly in His-Vif refolding.

A stepwise dialysis procedure was used to refold His-Vif protein. Briefly, 10 mL of purified His-Vif was diluted with chromatography buffer A to approximately 1 mg/mL. Denaturant – GnCl was slowly removed by a series of dialyses with buffers of decreasing GnCl concentration. The GnCl concentration in a base buffer of 50 mM NaCl, 20 mM HEPES, 5mM DTT and 5 mM EDTA, 5% glycerol, pH 7.4 was reduced as follows: 6 M→ 3 M→ 2 M→ 1

M→ 0.5 M→ 0 M of GnCl. After the final overnight dialysis with the base buffer without GnCl, the sample was removed from the dialysis tubing. The precipitated protein was removed by centrifugation at 18,000 rpm for 30 min at 4 °C. Purity of the refolded His-Vif was examined by SDS–PAGE with Coomassie blue staining. Protein concentration was measured using the BCA Protein Assay Kit (Pierce) according to the manufacturer's instruction.

Western blot analysis

Samples containing His-Vif protein were loaded onto 4-20% gradient SDS-PAGE (BIO-RAD). Following electrophoresis, proteins were transferred onto PVDF membrane using a Criterion™ Blotter (BIO-RAD) and blocked with 5% BSA in PBS with 0.2% Tween 20 (PBST). Then, the membranes were incubated with either a 1/1,000 dilution of His-probe, or anti-HIV-1 Vif (Santa Cruz Biotechnology) for 1h at room temperature. After washing with PBST three times for 10 min each, the membranes were incubated with an alkaline phosphatase-coupled rabbit anti-mouse IgG antibody (1/5,000 diluted; Pierce) for 30 min and washed again. The presence of recombinant Vif was detected by NBT/BCIP Substrate (Sigma) according to manufacturer's instruction.

Circular dichroism spectroscopy

Circular dichroism (CD) spectra of recombinant Vif protein was recorded and analyzed to assess its secondary structures. The measurements were performed with a Jasco J-715 CD spectrometer using a 0.1 cm path length cell. Temperature was maintained at 25 °C. Each sample contained 10 μM Vif (2.5 mg/mL) in 10 mM potassium phosphate, pH 7.4. Scans from 190 –

260 nm were recorded in 1 nm increments. Calculations of secondary structure contributions based on the CD spectra were performed using the CDSSTR algorithm [22] and the DICHROWEB server [23].

NMR spectroscopy

NMR spectra were recorded with Oxford spectrometers operating at 600 MHz (^{1}H). Sample temperatures were maintained at 25 °C. Gradient sensitivity-enhanced ^{1}H, ^{15}N-HSQC spectra were also collected. Samples containing ~0.3 mM Vif in 10 mM potassium phosphate, 50 mM NaCl, pH 7.4. Felix (Accelrys) was used for data processing and analysis. Na^+DSS^- in D_2O (0.00 ppm) was used for referencing the ^{1}H chemical shifts.

Thermal shift assay

Thermal shift assays were performed using an iCycler iQ Real Time PCR Detection System (Bio-Rad) in a total well volume of 25 µL containing 7.5 µl of 300 × Sypro Orange (Molecular Probes), 12.5 µl of 2 × test compound (compound library kindly provided by Dr. ZhiJie Liu, IBP, Beijing, China), and 5 µl of 2.5 mg/mL His-Vif were added to the wells of a 96-well thin-wall PCR plate (Bio-Rad). Buffer instead of test compound was added in the control samples. The 96-well microplate was heated from 25.0 to 95.0 °C with an increments of 0.5 °C/min. Fluorescence changes in the wells of the plate were monitored simultaneously with a charge-coupled device (CCD) camera. The wavelengths for excitation and emission were 490 and 575 nm, respectively. Protein thermal unfolding curves were monitored by detection of changes in fluorescence of the Sypro Orange. Melt curve data were analyzed using the custom

iCycler iQ Real Time PCR software. The melting temperature (Tm) was determined by reading the temperature at which the reaction was half- complete.

SAXS measurement

SAXS experiments for His-Vif were conducted in Rigaku Americas (Woodlands, Texas). SAXS data was generated using the Rigaku S-MAX 3000 BioSAXS system. The data was collected on the 200 mm multi-wire 2D detector at a distance of 1.5 m. SAXS measurements were carried out with the samples containing 8 mg/mL His-Vif in 20mM HEPES, pH 7.4, 150mM NaCl. For each protein sample analyzed, a buffer blank was also recorded (same exposure time) for buffer subtraction. Each sample was tested in duplicates using 30 and 60 minute exposure time, and the data were processed using SAXSGUI, ATSAS softwares. The scattering profile was determined as the difference between samples with and without proteins. Values for maximum pair distances were extracted using the program GNOM [24]. The P(r) output data was then used to calculate electron-density envelope. Each envelope is the product of 10 GASBOR [25] runs averaged with DAMMAVER [26].

RESULTS AND DISCUSSION

Cloning and expression of Vif in *E. coli*

To generate the expression vector – pDEST527-His-Vif, PCR-amplified *vif* cDNA was cloned downstream of 6×His tag by the Gateway recombinant cloning procedure. In order to express the Vif fusion protein, the plasmid pDEST527-His-Vif was transformed in *E. coli* Rosetta 2 competent cells. Initial SDS-PAGE analysis showed that His-tagged Vif protein was

expressed in inclusion bodies as shown in Figure 2.2. Several expression parameters, including post-induction temperature (18 or 37°C), time of induction (1-4h for expression at 37°C, overnight for expression at 18°C), the concentration of IPTG (0.2–1 mM), and different *E. coli* expression strains were evaluated in order to increase protein solubility. However, none of the conditions used showed a significant improvement in His-Vif solubility (data not shown). We also constructed GST- and MBP-tagged Vif for *E. coli* expression and these two fusion proteins were aggregated in the inclusion bodies as well. In all cases, the analysis of the insoluble fraction of His-Vif by SDS–PAGE showed a protein band at approximately 25 kDa corresponding to the fusion of 6×His (1 kDa) and full-length HIV-1 Vif (23 kDa). The identity of this band was confirmed to be His-Vif by western blot using the monoclonal anti-6×His and anti-HIV-1 Vif antibodies, as shown in Figure 2.2.

Purification and refolding of His-Vif

Inclusion bodies containing His-Vif in the cell lysate were pelleted by centrifugation and washed with urea and Triton X-100 to remove *E. coli* membrane and cell wall material. The urea concentration selected in the wash buffer is 2 M, because higher concentrations result in partial solubilization of the recombinant protein. A final wash step with the buffer containing no Triton X-100 ensures to remove excessive detergent from the inclusion bodies. GnCl appeared to be more efficient in solubilizing His-Vif than urea in this study (data not shown). An extensive purification of GnCl solubilized His-Vif was achieved using nickel affinity columns as shown from the chromatogram and the SDS-PAGE analysis (Figure 2.2 d). A single peak (blue line) about ~1150 mAU (OD_{280}) was eluted at approximately 250 mM imidazole in the gradient elution profile. Basically, 10 mg denatured His-Vif were produced from one gram bacterial cell

pellet. The SDS–PAGE (Figure 2.2 d insert) showed the purified His-Vif in a pool of the peak elutes stained with Coomassie blue. The purity of the protein from this purification step is estimated to be ~90%, and the eluted His-Vif was confirmed by Western blot using the anti-6×His and anti-Vif antibodies.

We also expressed ^{15}N-labeled His-Vif in *E. coli* cells grown in M9 media containing ^{15}N-labeled NH_4Cl as the sole Nitrogen source, as described in the experimental methods. Since the growth of *E.coli* was slower in M9 media than in LB media, a longer growth time (16 h) was employed. Purification was performed by the same procedures as described for the unlabeled His-Vif. A similar purity and yield were achieved.

Since the His-Vif from inclusion bodies was purified in 6 M GnCl as a denatured protein, it has to be refolded for further structural or functional studies. The refolding of His-Vif was accomplished by a stepwise dialysis with decreasing GnCl concentration from 6 M to 0 M as described in Experimental methods. The overall yield of refolded His-Vif was approximately 10 mg/L of bacterial culture.

Our results also revealed that refolded full-length Vif expressed from *E.coli* had a strong tendency to form multimers. Specifically, refolded Vif fusion protein was directly loaded onto a 4-20% Tris/glycine SDS-PAGE for electrophoresis, boiling with Laemmli sample buffers containing 2% SDS and 5% β-mercaptoethanol (BME) for 5 min. Under this denaturing condition, His-Vif protein migrated in the gel as monomers (25 kDa) and dimers (50 kDa) as shown in Figure 2.2.e. The majority of His-Vif was in the form of monomers. Similar results have been reported that, like other proteins encoded by HIV-1, Vif has a strong tendency toward self-association through its C-terminal PPLP domain [27]. Studies using dynamic light scattering, circular dichroism and fluorescence spectroscopy demonstrated that wild-type Vif

formed oligomers of five to nine proteins *in vitro*, while PPLP domain mutated Vif still formed dimers and/or trimers [28]. Therefore, the multimerization of Vif is not likely to be conclusively sequence-dependent. Thus far, the biological significance and mechanism of Vif multimerization are not well-studied, however *increasingly efforts* have been focused on this feature of Vif, since the multimerization is directly related to the Vif-mediated A3G degradation and the potential antiviral drug target [29, 30]. In order to investigate the mechanism of Vif multimerization, we also expressed and purified a truncated $Vif_{89\text{-}192}$ in which the Vif N-terminal A3G binding region was deleted. Surprisingly, the truncated Vif could form a series of higher-order multimers up to 10-mers or higher, comparing to dimers and trimers formed by full-length Vif (see Chapter 3 for details). Our results indicate that not only the Vif C-terminal PPLP motif, but also the residues from Vif N-terminal region were involved in Vif multimerization, possibly through a negative cooperative binding manner. In solution, Vif exists as complexes consisting of oligomers less than 8 to 10-mers [28]. Little is known on the relationship between Vif multimerization and its medicated A3G degradation, but either the deletion or mutation of the proposed PPLP Vif multimerization motif is able to severely decrease the Vif function in the viral life cycle [27, 28]. The mutations of PPLP motif not only reduce the Vif binding to A3G [31], but also diminish the binding to A3G mRNAs and the further inhibition of A3G translation [32]. However, there is no direct experimental evidence whether the enhanced Vif multimerization could increase its affinity to A3G and A3G mRNAs, and therefore to enhance A3G degradation. Therefore, further elucidation of the correlation between Vif multimerization and A3G degradation is of great importance. Based on our results, we suggest that the binding of A3G to Vif N-terminal region promotes the Vif self-association and possibly increases Vif affinity to both A3G and A3G mRNAs. This A3G-induced formation of Vif higher-order multimers, in turn, enhance the

binding capability for more A3G molecules, which help explain how Vif can effectively mediate the formation and degradation of high molecular mass (HMM) complex of A3G [33].

Circular dichroism studies

The secondary structure contents of the refolded Vif protein were determined using UV-CD spectroscopy. The CD spectrum of Vif protein in a physiologic neutral pH buffer at room temperature shows a minimum at ~202–203 nm giving a value of –16,259° cm^2/dmol $residue^{-1}$ indicating that the full-length Vif is largely unstructured as illustrated in Figure 2.3. The Analysis of the UV-CD spectrum for Vif revealed that it contains 14% α-helix, 36% β-sheet, and 50% unstructured random coil. The disordered structure of a chemically synthesized peptide C-terminal $Vif_{141-192}$ domain has been recently demonstrated using CD and NMR measurements [18]. In addition, the mass spectrometry analysis of chemically cross-linked *E. coli*-expressed recombinant Vif showed that Vif C-terminal region is mostly unstructured and N-terminal region is more protected and structured [21]. Consistent with the disorder sequence predictions, our results showed that most of the structured β-sheet elements are located at the Vif N-terminal A3G binding region and the unstructured regions interspersed by short α-helices are at the C-terminal portion. Overall, without binding with other proteins, Vif may exist as an unstructured protein.

1D and 2D-homonuclear NMR

To detect the presence of the tertiary structure of the purified Vif, we measured the one-dimensional proton NMR spectrum in solution conditions that resemble the physiological environment. For a well-folded protein, the 1H NMR signals should be well dispersed because

each proton group experiences a unique chemical environment. In contrast, as shown in Figure 2.4 a, the ^{1}H spectrum of Vif shows no signal dispersion typical of well-structured proteins, the 1D proton NMR signals of Vif in this study were not well dispersed. The spectrum showed narrow, intense and overlapped methyl proton resonances in the random-coil conformation within the values of 0.8 to 1.0 ppm, whereas resonances in the methyl regions of folded proteins are more spread out since they are more orderly packed. In addition, the backbone amide protons signals of Vif are rather low and are almost entirely clustered in a narrow region of ~7.6 to 8.4 ppm, whereas the amide resonances for known folded proteins are dispersed from ~7.5 to 10.2 ppm [34]. Thus, the NMR features of Vif, especially the lack of signal dispersion, represent a strong indication that Vif polypeptide chain lacks a stable defined tertiary structure in solution, which is consistent with the unfolded conformation of the Vif based on its CD spectrum.

We also performed the $^{1}H–^{15}N$ HSQC spectrum of Vif. The results from 2D spectrum suggest that the Vif protein is predominantly in a homogeneous state under the conditions used, as shown in Figure 2.4.b. Limited chemical shift dispersion is observed in the spectrum, as the majority of peaks appear between 7.8 and 8.5 ppm on the ^{1}H axis. The narrow range of peak positions is consistent with the characteristic of unstructured proteins [35]. Random coil position of side chain NH_2 groups from Asn and Gln residues produce a pair of ^{1}H peaks within 6.8 to 7.6 ppm corresponding to a single ^{15}N value, which are observed at approximately 112 ppm. In the spectrum of Vif, all the side chain amides have nearly equivalent chemical shifts to each other, suggesting they are equally solution exposed. In addition, the 2D Vif spectrum lacks the resonance peaks characteristic of higher-order structure which can neither be used to obtain complete side chain and backbone resonances assignment, nor distinguish regions of α-helix

from random coil. But the spectrum does indicate that Vif exists largely in an unstructured conformation without distinct tertiary structure.

Thermal shift assay

Thermal shift assays were also performed on the recombinant Vif. In this analysis, Vif displayed atypical thermal denaturation curve characterized by the high fluorescence intensities at low temperature (increased by 1.6 fold) and a significantly decreased Tm value (51.9°C) compared to the BSA control (Tm = 56.0 °C), as shown in Figure 2.5. Since we know that the thermal shift is a sensitive method to monitor protein thermal unfolding process using an environmentally sensitive fluorescent dye. The dye molecule is quenched in aqueous solution but preferentially binds to the exposed hydrophobic interior of an unfolded protein at the unfolding temperature resulting in a readily detectable signal change [36]. A well-folded protein such as BSA exhibits a sharp increase in fluorescence signal between 55 - 65°C indicating a state transition from folded to unfolded. Whereas, Vif showed a relatively high level of fluorescence signal of approximately 4,400 at low temperatures, i.e. in the 20 - 30°C range preceding the initial unfolding phase, indicating the exposure of hydrophobic residues. Moreover, the signal offset in Vif unfolding phase (~46 – 55 °C) is around 500, which is 8-fold less than that of BSA (with a signal offset of ~4000). We also tested whether a folded structure could be induced in the purified Vif by screening a compound library consisting of a wide variety of approximately 500 compounds and additives. We found that all conditions yielded samples with high initial dye binding at low temperature, and no conditions yielded the typical sigmoidal unfolding curve of a well-folded protein (data not shown).

SAXS measurement

In addition to that the NMR spectrum failed to provide any information for Vif structure determination, trials of Vif crystallization for X-ray crystallography had no success (data not shown). As a result, we employed SAXS to study the solution structure of Vif. Although SAXS cannot provide the atomic-level resolution as X-ray crystallography does, it uses X-ray radiation to reveals dynamic conformations and assemblies of macromolecules. The SAXS data is transformed from reciprocal space to real space to generate the P(r)-distribution, which is a histogram of the interatomic vectors within the macromolecule. The P(r)-distribution (pair distance distribution) gives an estimate of the maximum dimension (dmax) of the particles in solution [37]. The SAXS profile represents the simultaneous scattering measurement of the macromolecule in all orientations [38]. In addition to the dimension, SAXS experiments can also calculate the radius of gyration (Rg) of the protein molecules in solution. Since an increase in Rg is generally consistent with an opening of the macromolecule whereas a decrease in Rg suggests compaction [39]. The R_g value is therefore an indication whether the protein is "folded" or "unfolded". Besides, SAXS data can be used to determine the volume and shape of the scattering particle to generate an electron density envelope structure.

The electron envelopes and the pair distribution plot of Vif are shown in Figure 2.6. In solution, His-Vif existed as a particle up to 20 nm (D_{max} = 200 Å) with a real space radius of gyration value of 53.99 ± 0.147. The large particle likely represents the multimerization of Vif. By fitting with proteins of similar MW, it is estimated that this particle can accommodate 8-10 Vif molecules, which is in agreement with the observation of 5 to 9 molecules obtained from dynamic light scattering measurement [28]. According to the equation $R_g = R_0 N^v$ proposed by

Kohn *et al.* [40], the estimated radius of gyration of a disordered protein with similar residue number is about 44.2, which is much smaller than the experimental data we obtained. One possible explanation is that Vif forms stable multimers in a state like a transient extended premolten globule in solution. Surprisingly, the Kratky plot (Figure 2.6.c), which can be used to evaluate the protein folding state (open or compact), reveals a prominent peak that is typically characteristic of folded proteins. The Kratky plot of a folded globular protein presents a typical bell shape with a clear maximum peak, whereas for a completely unfolded protein, instead of the maximum peak, the curve displays a plateau. In this study, the disagreement between SAXS and other biophysical experiments, such as CD and NMR that indicate Vif is an unstructured protein, might result from the fact that unbound Vif can form stable multimers, rather than random aggregation, and behaves similar to partially folded proteins. We propose that Vif, in solutions, can form stable complex by self-association. The core of the complex is very compact and consists of Vif C-terminal regions, while leaving its flexible N-terminal motifs free to interact with A3G. Therefore, the SAXS experiments on Vif-A3G complex and/or Vif-A3G-E3 ubiquitin ligase complex are of great importance to validate the bound and unbound Vif in-solution conformation.

CONCLUSION

HIV-1 Vif directly interacts with the host antiviral factor – APOBEC3G and leads it for degradation through the proteasome pathway. The A3G degradation also requires the binding of Vif with the cellular E3 ubiquitin ligase complex, majorly comprising cullin5, elongin B, and elongin C. The elongin C directly binds to Vif C-terminal SOCS BC-Box motif [41], while

Cullin5 binds to Vif HCCH motif [21, 42]. On Vif-mediated A3G degradation, Vif therefore is serving as an adaptor molecule to recruit the E3 ubiquitin ligase complex to its substrate – A3G, then the E3 ubiquitin ligase complex induces the polyubiquitination of A3G and directs it to the 26S proteasome for degradation. As a result, this function of HIV-1 Vif makes it a critical factor in HIV-1 infectivity and virus spread, which provides a potential pharmacological target.

In this study, we reported a structural analysis of the full-length HIV-1 Vif protein using various biophysical methodologies and provided direct evidence for the intrinsically unstructured nature of Vif under physiological conditions. However, the difficulty of expression and purification of full-length soluble Vif in prokaryotic and eukaryotic systems with a manageable quantity for structural and functional studies has greatly impeded the understanding of Vif functions. Instead of using truncated Vif domains or chemical synthesized peptides as most other groups previously reported, we are able to purify and refold the full-length Vif with a satisfactory quantity and purity. Our results revealed that Vif lacked substantial secondary structure elements under physiological conditions as indicated by CD and NMR results. Vif, by itself is an intrinsically unstructured protein that has regions with poorly defined tertiary structure when studied in solution. However, the Vif C-terminal region could possibly become more compact and less flexible upon multimerization as indicated by SAXS results. Based on the studies on unbound Vif, we strongly suggest that binding with its biological partners such as A3G and E3 ubiquitin ligase would make Vif more ordered. Because of the multimerization and intrinsically unstructured nature of Vif, both X-ray crystallography and NMR are not very likely to be the potent structure determination methodologies to elucidate the structure of unbound Vif. Therefore, structure determination using Vif complexed with other proteins including A3G,

Cul5, elongin C and antibodies, would be a reasonable way to obtain a high resolution atom-level structure.

During gel filtration and other chromatography experiments such as ion exchange chromatography in non-denatured conditions, it is difficult to elute full-length Vif from the columns, most likely not only due to the aggregation of Vif, but also some unknown Vif interactions with polysaccharides. Of note, this interaction was not interfered by the addition of detergents, salt and solvent such as DMSO.

Based on our results from extensive studies of HIV-1 Vif molecular structure using biophysical methods, we propose the following hypothesis: HIV-1 Vif monomers are likely in dynamic equilibrium in a low molecular weight homo-multimers (up to 8- to 10-mers). Vif N-terminal regions that may have a transiently defined topology interact with Vif C-terminus that is intrinsically disordered and associated with other Vif counterparts. Upon binding with biological macromolecules such as A3G or E3 ligase complex, the interaction between N- and C-terminus becomes diminished and the other binding region is released for binding with its partner proteins. Specifically, if A3G is bound to Vif, a higher molecular weight of Vif complexes would be temporarily formed, which helps to recruit more E3 ubiquitin ligase complex for degradation of the high molecular weight complex of A3G. Further structure determination and biochemical analysis of these complexes must be performed in order to better elucidate the detailed structure and functions of Vif. Future drug design involved in disruption of one or more of Vif interactions would be beneficial from the understanding of Vif complex structure.

Figure 2.1. The functional domains of HIV-1 Vif. Known HIV-1 Vif functional domains are aligned with the PONDR VL-XT prediction of Vif order/disorder regions [43-45]. Interaction partners are described along the top of the figure. The gray line in the PONDR plot represents the PONDR score of 0.5: regions with scores below 0.5 are likely to be structured, while a score above 0.5 indicates unstructured regions. The amino acid residue numbers are at the bottom of the plot, which is in accordance with the relative position of each major Vif function domains. The structural information therein is adapted from [20, 46]

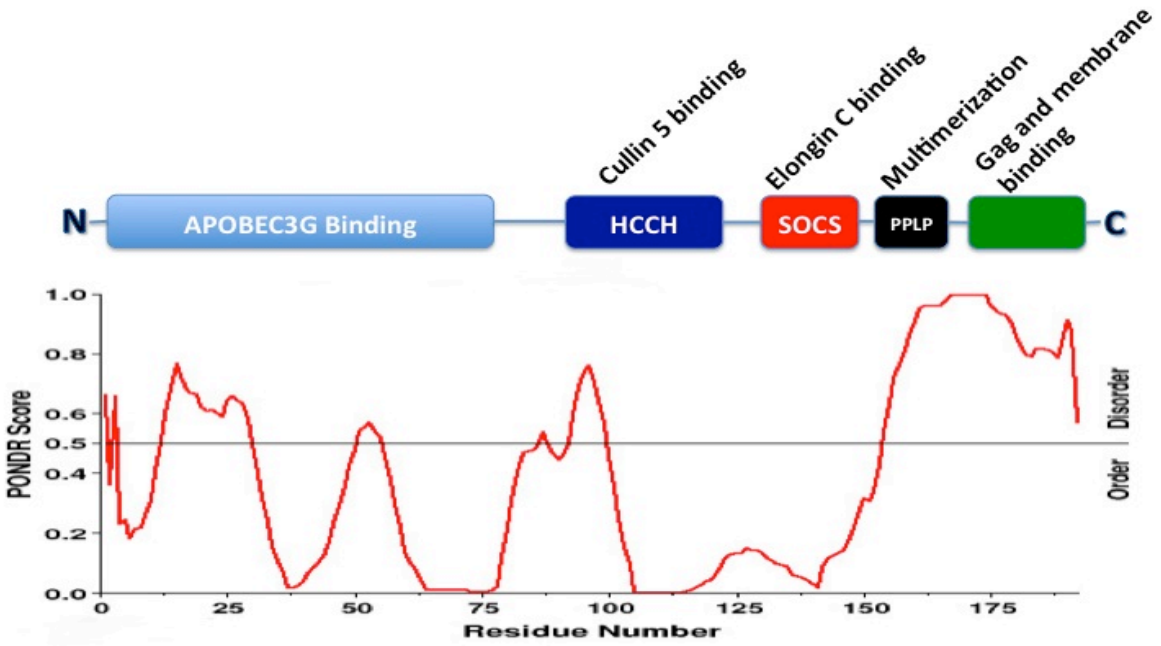


Figure 2.2. Bacterial expression and purification of recombinant Vif.

a, SDS–PAGE analysis of recombinant Vif expression in *E.coli*, uninduced cells (1) and IPTG induced cells (2); recombinant Vif was over-expressed with the MW of 25 kDa.

b, Western blot of uninduced cells (1) and induced cells (2) detected by anti-Vif antibody.

c, solubility assay of *E.coli* expressed Vif; bacterial cells were lysed and the insoluble fraction (lane P) was separated from the soluble fraction (land S) by centrifugation, the 25 kDa recombinant Vif was apparently present in the insoluble fraction of cell lysate. Lane I represents the IPTG induced cells.

d, chromatogram of HisTrap column purification of His-Vif in denaturing condition (blue line). The His-tagged Vif was eluted by an imidazole gradient (20–500 mM, green line); insert shows the SDS-PAGE analysis of elute fractions containing purified Vif.

e, SDS-PAGE analysis of refolded Vif showing Vif monomer (25 kDa) and Vif dimer (50 kDa) as indicated by arrows.

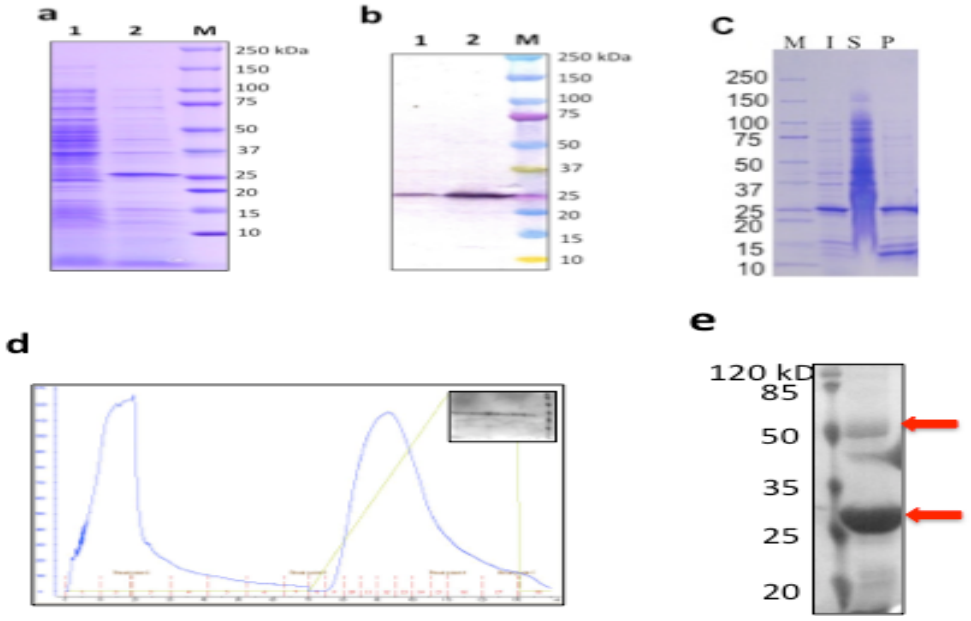


Figure 2.3. Circular dichroism spectrum for His-Vif in 10 mM potassium phosphate, pH 7.4 at room temperature. Protein concentration was 2.5 mg/mL. Five scans were averaged and subtracted from a buffer blank scanned under the same conditions.

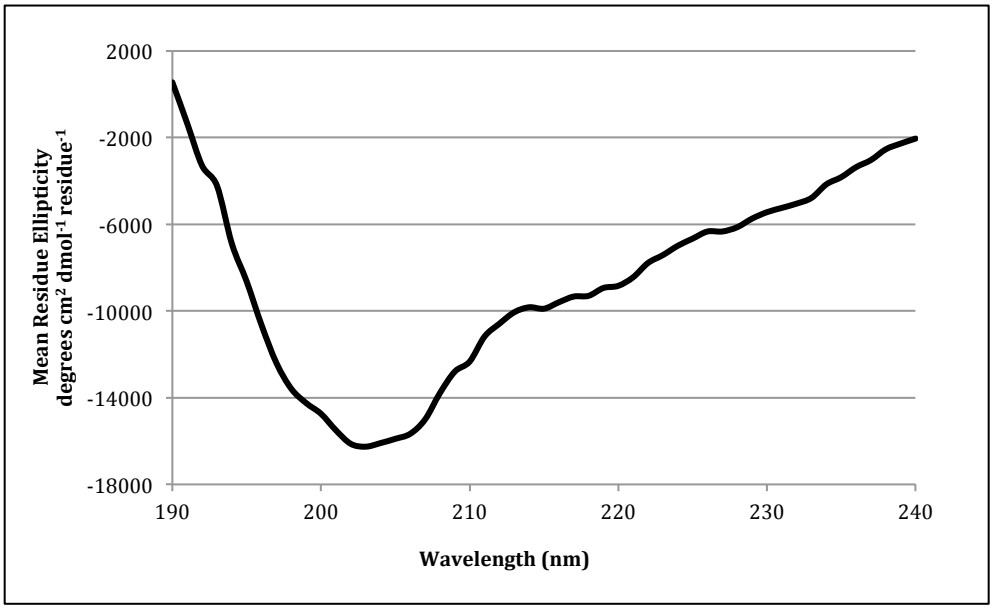


Figure 2.4. NMR characterization of the Vif. a, the one-dimensional ^{1}H NMR spectrum of the 300 μM protein was measured at 600 MHz in 10 mM potassium phosphate, 50 mM NaCl (pH 7.4), at 25 °C. b, $^{1}H–^{15}N$ HSQC spectrum of Vif, showing characteristics of a unstructured protein.

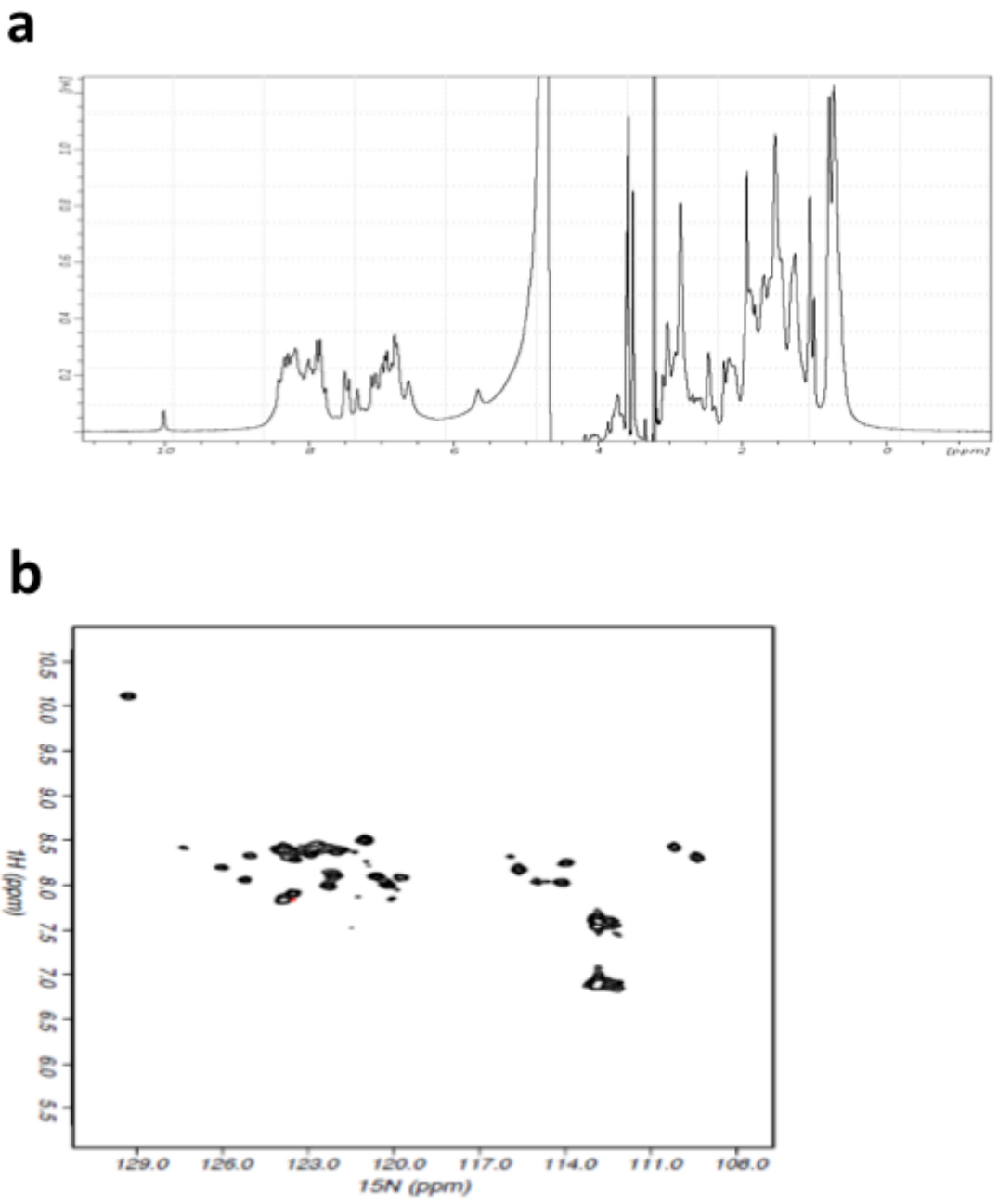


Figure 2.5. Thermal shift assay of Vif and BSA control. Unfolding of Vif (Red line) and BSA control (Blue line) monitored in the presence of SYPRO orange. The proteins were gradually heated and the unfolding process was monitored by detection of the charges of fluorescence of the dye reporter that binds preferentially to the unfolded proteins. The fluorescence intensities (Y axis) are measured according to the temperature (X axis, in °C).

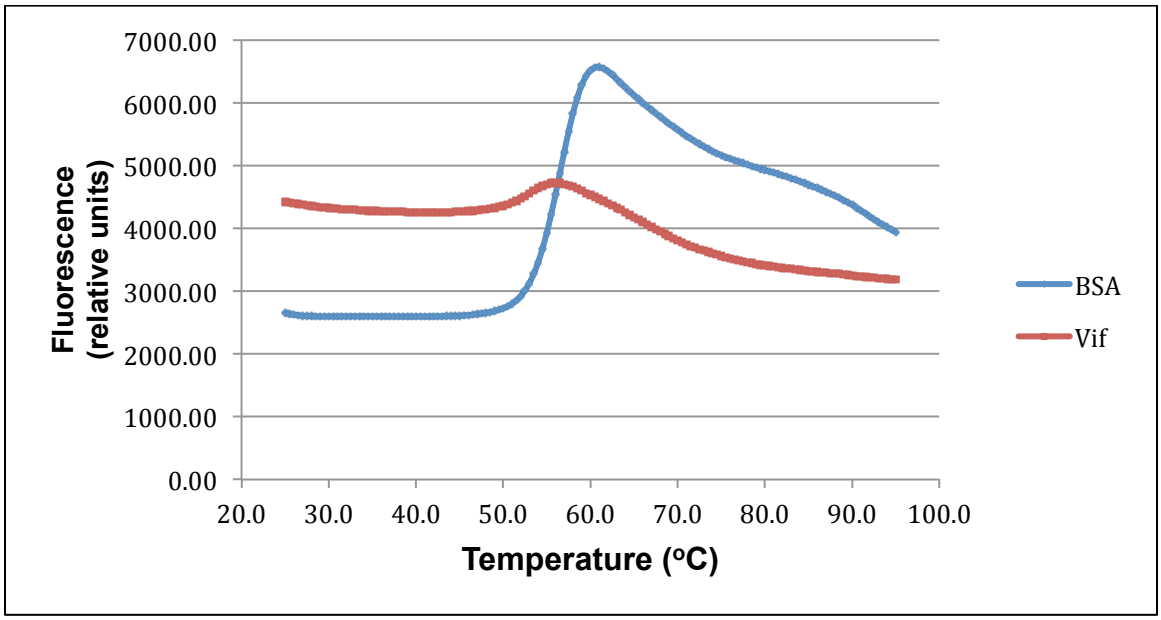
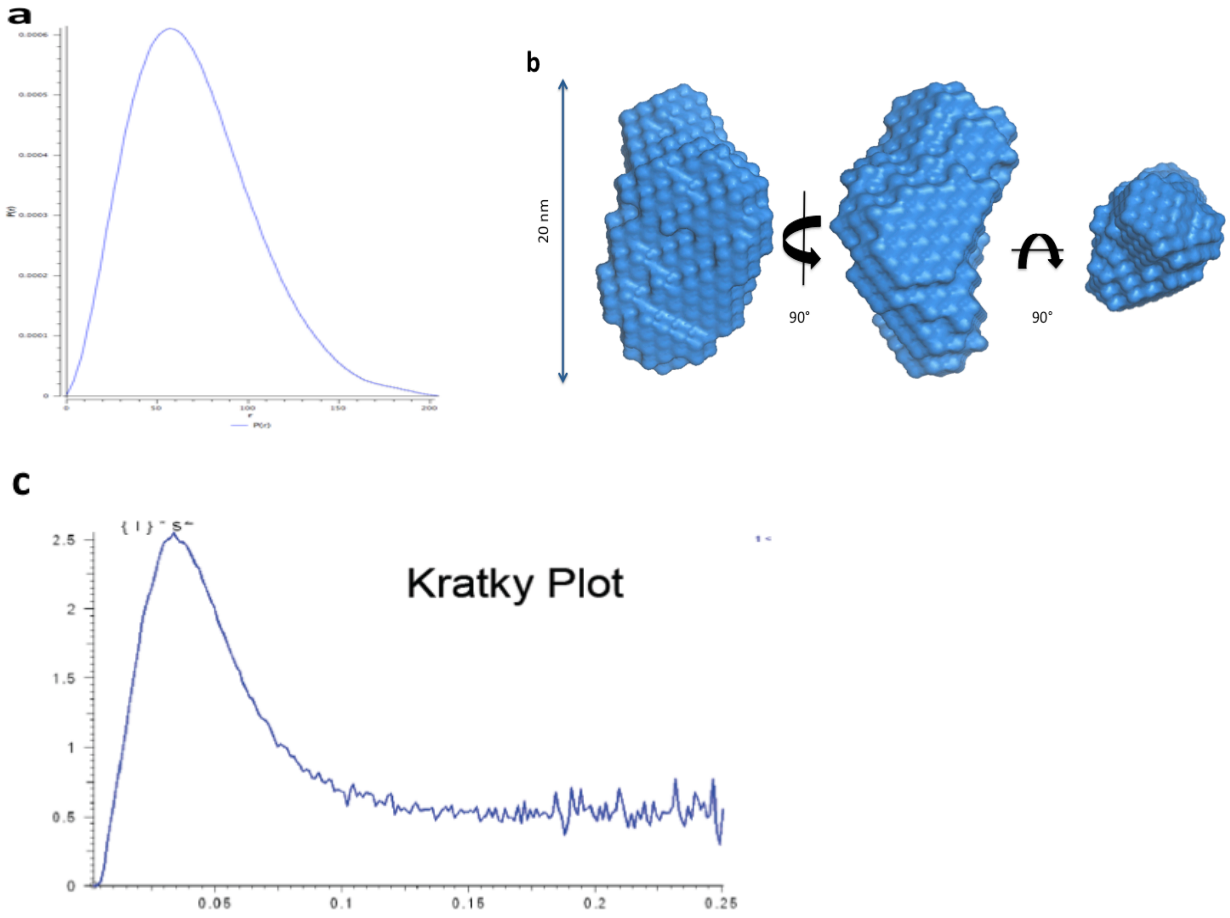


Figure 2.6: SAXS measurements of Vif in solution. a, pair distribution plot for Vif sample showing one large peak that has the D_{max} of 20nm. b, electron density envelope model of Vif (from left to right: front view, side view and top view) that shows the 20 nm Vif particle present in solution. c, the Kratky plot of Vif shows a typical bell shape with a clear maximum peak.



CHAPTER 3

BIOCHEMICAL AND BIOPHYSICAL INVESTIGATIONS OF THE INTERACTION BETWEEN HIV-1 VIF AND HUMAN APOBEC3G[2]

[2]Zhou, D., Xu, H., Wang, B.C., Liu., Z. and Rose, J.
To be submitted to *The Journal of Biological Chemistry*

ABSTRACT

The human immunodeficiency virus type-1 (HIV-1) Virus Infectivity Factor (Vif) mediates the degradation of a cellular antiviral factor, Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G (APOBEC3G or A3G), by serving as an adaptor that bridges the A3G and E3 ubiquitin ligase complex. The human A3G, a cytidine deaminase, is a potent inhibitor of retroviruses and transposable elements and functions to deaminate cytidines to uridines in HIV-1 ssDNA during viral replication. A3G contains two cytidine deaminase domains of which only the C-terminal one is known to have DNA deamination activity. This uneven distribution of enzymatic activity may suggest a dimeric conformation of A3G. In this study, we found that baculovirus expressed A3G is in monomeric and dimeric conformation that provides direct evidence for the dimerization of A3G. However, no higher-order A3G species such as trimer or tetramer were observed. By exploiting the Biacore surface plasmon resonance (SPR), the binding kinetics of Vif-A3G interaction was measured. Surprisingly, full-length Vif preferably interacts with A3G dimers (K_D=0.2 nM) with an affinity that is 1000 fold greater than that with A3G monomers (K_D=200 nM). Also, we confirmed that A3G interacts with the N-terminal region of Vif by comparing the interaction between A3G and truncated Vif_{89-192}. In addition, we observed that both purified full-length Vif and truncated Vif_{89-192} could form multimers, while Vif_{89-192} can form much higher-order multimers that were very stable in the presence of detergent, reducing agent, denaturing agent and heat. Accordingly, we suggest that the N-terminal A3G binding region of Vif may also be involved in Vif multimerization.

INTRODUCTION

HIV-1 Vif is a 23 kDa basic phosphoprotein, expressed late in the HIV-1 life cycle. Current knowledge indicates that Vif protein has no established enzymatic activity. Instead, it appears to serve primarily, if not exclusively, as an adaptor molecule to mediate the interaction of other viral or host factors. Initially, Vif was demonstrated to be required for production of infectious progeny virions in some, but not all cell types. In the absence of Vif, virions were estimated to be approximately 1000 times less infectious than the ones produced from wild-type virus [1]. Therefore, the function of Vif is cell species specific and is tightly linked to the nature of the virus-producing cell. Several T-cell lines (e.g. Jurkat, CEM-SS and SupT1) and non-haemopoietic cell lines (e.g. HeLa, 293T and COS) produce infectious HIV-1 virions in the absence of Vif (Δvif HIV-1) and are termed 'permissive'. Conversely, Δvif HIV-1 virions derived from nonpermissive cells, including primary CD4 T cells (e.g. HUT78, CEM) and macrophages, are non-infectious [2]. In other words, Vif is critical for virus replication in primary CD4 T cells and macrophages – the natural targets of HIV-1, yet entirely dispensable in other T-cell lines. Despite the early identification of permissive and nonpermissive cell types by supporting or not supporting the spread of Δvif HIV-1 respectively, the molecular basis for this implausible "Vif phenotype" remained elusive for many years. Also, the mechanism by which Vif allows wild-type HIV-1 virions to spread readily in nonpermissive cells remained unclear. Recently, Sheehy *et al.* succeeded in identifying human APOBEC3G as the host Vif-sensitive antiviral factor by screening two genetically similar cell lines that are showing opposite Vif phenotypes [3]. A3G, belonging to a family of cytidine deaminases, catalyzes hydrolytic deamination at the C4 position of the cytidine (or dC) base, converting cytidine to uridine (or dC to dU) on single-stranded DNA or RNA. During HIV-1 reverse transcription, A3G selectively

recognizes newly reverse-transcribed single-stranded viral DNA (first strand of viral DNA) and extensively catalyzes dC-to-dU deamination. The resulting G-to-A hypermutation is estimated to exceed 10% of all dG nucleotides in HIV-1 DNA, which is excessive and sufficient to inhibit the viral infection and replication [4]. Although A3G can effectively cause extensive mutations on HIV-1 DNA due to its cytidine deaminase activity, recent studies have revealed that A3G can also exert deaminase-independent antiviral activity [5, 6]. A3G can interfere with primer tRNA annealing, minus- and plus-strand DNA transfer, primer tRNA progression, provirus formation, and viral DNA elongation and integration, among them the primary causes for A3G-caused HIV-1 reduction are that A3G inhibits viral cDNA synthesis by impeding RT elongation and interfering viral DNA integration [7, 8]. However, A3G alone can effectively inhibit the replication of ΔVif HIV-1 but not that of wild type HIV-1 due to the fact that the major function of Vif is to counteract A3G by inducing the polyubiquitylation and subsequent proteasome degradation of A3G. To degrade A3G, Vif hijacks a cellular E3 ubiquitin ligase complex consisting of cullin5 (Cul5), elongin B, elongin C, and a ring finger protein (Rbx). Initially, Vif binds to A3G and E3 ubiquitin ligase complex, therefore serving as an adaptor molecule to recruit the ligase complex to its substrate – A3G; then the E3 ubiquitin ligase complex induces the polyubiquitination of A3G and directs it to the 26S proteasome for degradation [9]. Rapid advancement has been made in the studies of the biological functions of A3G in recent years. The structural and functional studies of the interaction between A3G and HIV-1 Vif have provided a great amount of information for understanding the mechanisms by which Vif mediates the antiviral activity of A3G. However, many questions remain unanswered mainly due to the lack of high-resolution three-dimensional structures of HIV-1 Vif. This lack of data is primarily because of the difficulty of expressing high levels of soluble recombinant Vif using

either prokaryotic or eukaryotic expression systems. Also, the purification of full-length Vif in a large amount is extremely difficult. Hence, Vif is the only HIV-1 protein that neither the crystal nor the solution structure is available. So far, the only available crystal structure of Vif is the α-helix composed of residues 140–155 within the Elongin B/C complex [10]. Therefore, any structural information either derived from modeling or indirect experiments would be very helpful to completely elucidate the function of Vif and its interaction with A3G. Overall, Vif consists of several functional domains as shown in Figure 3.1, including (1) the N-terminal RNA- and A3G-binding domains, (2) the central domain containing a His-Cys-Cys-His- zinc binding motif and Cul5 binding region, and (3) the C-terminal domain (CTD) that mediates the interactions with HIV-1 Gag protein, host cellular Elongin C and Vif-Vif multimerization. It has been reported that the C-terminal domain residues 141-192, which mediate interactions with both Elongin C and Cul5, are intrinsically disordered [11-13]. The mass spectrometry study on the chemically cross-linked Vif also indicated that the N-terminal region of Vif is likely to be protected or folded in a compact domain [14].

In this study, we provide direct evidence for the interaction between full-length HIV-1 Vif and A3G using biochemical and biophysical methods including pull-down assay, AlphaScreen and Biacore surface plasmon resonance. For the first time, A3G dimers were purified in this study which provides direct evidence for the dimerization of A3G. By comparing the binding kinetics of A3G monomer and dimer with Vif, we found that the A3G dimer has a binding affinity almost 1000 times greater than that of A3G monomer. Also, we confirmed that Vif N-terminal region accommodates the A3G binding sites. Two Vif proteins – full-length and truncated $Vif_{89\text{-}192}$ were prepared to study the A3G binding properties; however, they showed

distinct multimerization patterns that indicated that truncated Vif could form stable higher-order multimers.

EXPERIMENTAL METHODS

Construction of Vif fusion proteins *E. coli* expression vectors

Full-length HIV-1 Vif and truncated $Vif_{89\text{-}192}$ were constructed by PCR using pcDNA3.1 vector containing full-length HIV-1 Vif cDNA (a gift from Dr. Tom Hodge at University of Georgia) as the template. All the sequences of primers that were used in this study are given in Table 3.1. The PCR products were purified by agarose gel electrophoresis and were isolated using a QIAquick Gel Extraction kit (QIAGEN), followed by cloning into pDONR221 plasmid to generate an entry clone using BP Clonase (Invitrogen). The resulting plasmid, pDONR221-Vif and pDONR221- $Vif_{89\text{-}192}$ were used as the basis for expression plasmid construction. The *vif* genes were then cloned into the Gateway expression vector pDEST527 (N-terminal 6×His tag) by recombination reaction between the entry clone and the destination vector using LR Clonase (Invitrogen) to generate the expression clone pDEST527-His-Vif and pDEST527-His-$Vif_{89\text{-}192}$, respectively, according to manufacturer's instruction. Briefly, pDONR221-Vif and pDEST527 (150 ng each) were added to a 10 μL LR Clonase enzyme reaction mixture and incubated for 1 h at 25 °C to generate the destination vectors for protein expression. Next, *E. coli* TOP10 competent cells were transformed with 1 μL of each of the recombinant destination plasmid - pDEST527-His-Vif or pDEST527-His-$Vif_{89\text{-}192}$, and selected on Luria–Bertani (LB) agar plate containing 100 μg/mL ampicillin. Positive clones were identified by PCR, and confirmed by DNA sequencing.

Expression of His-Vif fusion protein

For expression of recombinant full-length and truncated HIV-1 Vif fusion proteins, *E. coli* Rosetta 2 Competent Cells (Novagen) were freshly transformed with expression vector pDEST527-Vif or pDEST527-Vif_{89-192}. Cells from a single colony were first grown in a small scale of 50 mL LB broth supplemented with 100 μg/mL ampicillin. Cells were cultured at 37 °C with shaking at 250 rpm. Next morning, 20 mL of the overnight culture were used to inoculate a 1 L culture. The expression of recombinant Vif proteins was induced by the addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside) to 1 mM when the cell density OD_{600} reaches 0.5. After grown for an additional 4 h at 37 °C, cells were harvested by centrifugation for 10 min at 5000 ×*g* at 4 °C and cell pellets were frozen at −80 °C until use. The solubility of recombinant Vif fusion proteins was analyzed using CelLytic B Plus Kit (Sigma-Aldrich) following manufacturer's instruction.

Purification and refolding of His-Vif fusion proteins

For isolation of inclusion bodies, 10g of frozen cell pellets expressing 6×His-tagged full-length Vif and truncated Vif_{89-192} were resuspended in 100 mL of ice cold PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4•2H_2O$, and 1.76 mM KH_2PO_4) containing 5 mM DTT (Dithiothreitol), 1 mM PMSF (phenylmethylsulfonyl fluoride), Complete Protease Inhibitor Cocktail (Roche), and 1 mg/mL lysozyme, pH 7.4. The lysis suspensions were incubated 30 min at room temperature with gentle agitation, followed by sonication on ice for 3 min. Lysates were cleared by centrifugation at 21,000 ×*g* (JA-25.50 Rotor in Beckman Coulter Avanti® J-E centrifuge) for 30 min at 4°C. The resulting pellets were resuspended with 50 mL of wash buffer

(50 mM Tris–HCl, pH 8.0, 2% Triton X-100 (v/v), 500 mM NaCl, 1 mM 2-mercaptoethanol, 1 mM PMSF, and 2 M urea) and centrifuged at 21,000 ×*g* for 20 min at 4°C. This step was repeated three times with 50 mL Triton X-100 free wash buffer for the final wash. Following the final centrifugation, the washed inclusion bodies were solubilized in 10 mL of solubilization buffer containing 6 M Guanidine–HCl (GnCl), 50 mM Tris–HCl, pH 8.0, and 1 mM 2-mercaptoethanol by incubation at room temperature for 6 h. The solutions containing denatured Vif fusion proteins were cleared from insoluble material by centrifugation at 21,000 ×*g* for 30 min at 4 °C, followed by filtration through 0.45 μm syringe filter.

His-tagged Vif proteins solubilized in 6 M GnCl was further purified with a 5 mL HisTrap™ HP Immobilized Metal Affinity Chromatography (IMAC) column connected in AKTA Prime system (Amersham–Pharmacia Biotech). The denatured protein samples were loaded onto the column pre-equilibrated with chromatography buffer A containing 6 M Guanidine–HCl, 50 mM Tris–HCl pH 8.0, 1 mM 2-mercaptoethanol, and 20 mM imidazole. After loading, the column was washed with 10 column volumes of buffer A. The His-tagged Vif fusion proteins were eluted in a 6-column volume linear gradient of 20–500 mM imidazole in buffer A. Pooled fractions containing Vif proteins were either stored at −80 °C or used directly in refolding.

A stepwise dialysis procedure was used to refold His-Vif proteins. Briefly, 10 mL of purified His-Vif was diluted with above mentioned chromatography buffer A to approximately 1 mg/mL. Denaturant – GnCl was slowly removed by a series of dialyses with buffers of decreasing GnCl concentration. The GnCl concentration in a base buffer of 50 mM NaCl, 20 mM HEPES, 5 mM DTT and 5 mM EDTA, 5% glycerol, pH 7.4 was reduced as follows: 6 M→ 3 M→ 2 M→ 1 M→ 0.5 M→ 0 M GnCl. After the final overnight dialysis with the base buffer

without GnCl, the samples were removed from the dialysis tubing. The precipitated proteins were removed by centrifugation at 18,000 rpm for 30 min at 4 °C. Purity of the refolded His-Vif and His-Vif$_{89\text{-}192}$ were examined by SDS–PAGE with Coomassie blue staining. Protein concentration was measured using the BCA Protein Assay Kit (Pierce) according to the manufacturer's instruction.

Construction of APOBEC3G fusion protein baculovirus expression vector

The plasmid *pcDNA3.1-hA3G*, which encodes hAPOBEC3G (A3G), was kindly provided by Dr. Chia-Kuei Wu. Full-length human APOBEC3G cDNA was amplified by PCR using A3G for and A3G rev primers containing EcoRI and XbaI restriction sites respectively, as listed in Table 3.1. The EcoRI/XbaI-digested A3G cDNA fragment was inserted into a modified donor vector in the Bac-to-Bac baculovirus expression system – *pFASTBAC-MBP* (containing N-terminal 6×His and MBP fusion tags in tandem, and a TEV protease recognition sequence; kindly provided by Dr. Chia-Kuei Wu) that had been digested with EcoRI and XbaI, resulting in pFASTBAC-MBP-A3G. The orientation and sequence of A3G insert were verified by DNA sequencing.

Virus production

The pFASTBAC-MBP-A3G plasmid described above was used to generate recombinant baculovirus carrying the desired A3G insert. Baculovirus vector was produced following the manufacturer's procedures described for the Bac-to-Bac system (Invitrogen). Briefly, the pFASTBAC-MBP-A3G plasmid was transformed into DH10Bac competent cells containing the baculovirus genome bacmid. After screening, white colonies containing recombinant bacmid

with the insert were selected. The recombinant bacmids were purified and analyzed by PCR for the presence of the A3G sequence.

Spodoptera frugiperda Sf9 cell maintenance

Sf9 cells adapted to HyClone SFX-Insect serum free medium (Thermo Scientific) supplemented with 1% Fetal bovine serum (FBS) and 1 × gentamicin were cultured at 27 °C in either monolayer or suspension culture. Sf9 suspension cultures were maintained at a cell density of 1×10^6 to 5×10^6 cells/mL as 50 mL cultures in 250 ml shaker flasks shaking at 95-120 rpm.

Insect cell transfection and recombinant virus isolation

To generate baculovirus for A3G intracellular production, newly attached Sf9 cells in 6-well plates (1×10^6 cells/well) were transfected with 1 μg of PCR-verified recombinant A3G bacmid DNA using Cellfectin II™ (Invitrogen) according to the manufacturer's instruction. The transfected Sf9 cells were incubated in 2 mL unsupplemented HyClone SFX-Insect serum free medium for 5 h at 27 °C. After removal of the transfection mixture, 2 mL SFX-Insect medium supplemented with 10% FBS and 1× gentamicin was added to each well. The medium that contains A3G baculovirus was collected 5 days post-transfection and reserved as P1 viral stock. A higher titer P2 viral stock was amplified in Sf9 suspension culture. Specifically, 50 mL of Sf9 cells at a cell density of 2×10^6 cells/mL was infected with 1 mL of P1 virus and the supernatant containing P2 virus was harvested 3 days post-infection. The resulting P2 viral stock was used for large-scale A3G expression.

Expression and purification of recombinant A3G

For large scale A3G intracellular expression, a typical Sf9 suspension culture (500 ml with 2.5×10^6 cells/mL) in a 2 L shaker flask maintained in supplemented SFX-Insect medium was infected with 500 μL of P2 virus from stock and cultured at 27 °C shaking incubator at 95 rpm. Infected cells were harvested 72h post-infection and recovered from medium by centrifugation at 2000*g* for 10 min. The pellet was washed twice with ice-cold PBS, collected by centrifugation and frozen at −80°C until use. The cell pellets containing A3G fusion protein were resuspended in 1/40 volume of original cell culture with ice-cold lysis buffer containing 10 mM phosphate buffer, pH 7.4, 5% v/v glycerol, 300 mM NaCl, 20 mM imidazole, 0.1 mM DTT, 0.1 mM EDTA, and protease inhibitor cocktail (Roche). Crude cell extracts were prepared by a treatment consisting of 5 passes through a French press operated at 4 °C. Lysates were cleared by centrifugation at 18,000 rpm (JA-25.50 Rotor in Beckman Coulter Avanti® J-E centrifuge) for 30 min at 4°C and the supernatant was subjected to affinity chromatography. Two types of affinity columns and one gel filtration column were used for purification including MBPTrap™ HP (GE Healthcare) for MBP tag, Ni Sepharose 6 Fast Flow resin, HisTrap™ HP IMAC column (GE Healthcare) for 6×His tag, and a Superdex 75 10/300 column (Figure 3.2). Specifically, the clarified, crude lysate was incubated with pre-equilibrated Ni Sepharose 6 Fast Flow resin at a ratio of 1 ml beads/25 mL lysate for 2 h at 4 °C. After binding, the resin was collected in an empty column and washed twice with 20 resin volumes of binding buffer containing 10 mM phosphate buffer, 5% v/v glycerol, 300 mM NaCl, 20 mM imidazole, 0.1 mM EDTA, and 0.1 mM DTT, pH 7.4, and eluted with ten bed volumes of binding buffer containing 500 mM imidazole. Immediately, the elution was further purified on a 5 mL MBPTrap™ HP column

connected in the AKTA Prime system (Amersham–Pharmacia Biotech). The sample was loaded onto the column pre-equilibrated with MBP chromatography *buffer A* containing 10 mM phosphate buffer, 5% v/v glycerol, 300 mM NaCl, 0.1 mM EDTA and 0.1 mM DTT, pH 7.4. After sample loading, the MBPTrap column was washed with ten column volumes of MBP buffer A. The A3G fusion protein was eluted with 20 mL of MBP chromatography buffer B containing 10 mM phosphate buffer, 5% v/v glycerol, 300 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT and 10 mM maltose, pH 7.4. Since the fusion tags and A3G were linked by a TEV protease recognition sequence, which can be cleaved by TEV protease. Partially purified A3G fusion protein eluted from MBPTrap column was incubated with TEV protease overnight at 16 °C with a ratio of 1 μg TEV protease/50 μg total protein. The completeness of cleavage was analyzed by SDS-PAGE and Western blot before loading onto the HisTrap™ HP column to separate the A3G from 6×His-MBP fusion tags. The cleaved A3G is present in the flow-through of the second HisTrap™ HP column purification, while the fusion tags bind to the column due to the presence of 6×His tag at the N-terminus of the construct.

Next, the A3G protein was concentrated to approximately 10 mg/mL and further purified by size exclusion chromatography at 15 °C using an AKTA Prime system (GE Healthcare) fitted with a Superdex 75 10/300 column. Buffer containing 20 mM HEPES, pH 7.4 containing 150 mM NaCl, 5% glycerol, 5 mM EDTA, 5 mM DTT was used as the mobile phase at a flow rate of 0.5 mL/min. The fractions containing purified A3G were aliquoted and stored frozen at −80 °C.

AlphaScreen studies

Analysis of direct binding of His-tagged full-length Vif protein to A3G was performed using an established AlphaScreen assay (Perkin Elmer) according to the manufacturer's protocol.

Briefly, reactions were conducted in 20μL final volume in triplicate in 384-well Optiwell™ microtiter plates (Perkin Elmer). The reaction buffer contained 20 mM HEPES, pH 7.4, 150 mM NaCl, 0.01% (v/v) Tween-20 and 0.1% (w/v) BSA. Initially, *different amounts* of His-Vif (10-150 ng) were incubated with Ni^{2+}-chelate donor beads (20 μg/mL final concentration) in a 10μL reaction volume for 1 h in the dark at room temperature. Then 10 μl of A3G-conjugated Acceptor beads were mixed with 10 μl of His-Vif Donor beads to the final concentration of 10 μg/mL of each bead. Reactions were incubated for 30 min at room temperature in 384-well plates in a volume of 20 μl. Binding interactions were then detected by measuring light emission signals in 10 min intervals in the EnVision reader (Perkin Elmer) and analyzed using the EnVision manager software. The unconjugated acceptor beads were used as controls.

His pull-down assay

The His pull-down assay was performed using the refolded full-length His-Vif fusion protein to pull down purified A3G. Briefly, 5 μg of refolded His-Vif fusion protein was bound to 100 μL Ni Sepharose resin in binding buffer containing 20 mM HEPES, pH 7.4 containing 150 mM NaCl, 5% glycerol, 0.1% Triton X-100 for 4 h with rotation. The resin was then collected in an empty column and washed twice with 10 mL binding buffer. 10 μg A3G diluted in binding buffer was then added to the column and incubated for 30 min at room temperature. The complex was washed with 10 mL binding buffer containing 20 mM imidazole to prevent nonspecific binding. The flow-through and the complex-bound resin were then boiled for 5 min in SDS-PAGE sample buffer and analyzed by SDS-PAGE. A3G was also incubated with Ni Sepharose beads alone as the negative control.

Biacore surface plasmon resonance studies

Biacore experiments were performed to determine the binding kinetics between the purified A3G and Vif proteins using Biacore 3000 instrument (GE Healthcare) at 25°C. The full-length His-Vif and His-Vif_{89-192} were covalently conjugated to the CM5 sensor chip respectively, as stationary phase, while A3G proteins were used as mobile phase in the experiment. A total of 1500 response units of either full-length His-Vif or His-Vif_{89-192} were immobilized via amine coupling to the chip flow chambers. Briefly, the surfaces of CM5 sensor chip flow chambers were activated by injecting 120 μL of 100 mm *N*-hydroxysuccinimide and 400 mm *N*-ethyl-*N*′-(dimethylaminopropyl)carbodiimide in equal volumes, followed by injecting 150 μL of 50 μg/mL each Vif fusion proteins diluted in 10 mm HEPES, 150 mm NaCl, 0.05% Tween-20, pH7.4 with the flow rate at 10 μL/min. The unreacted carboxymethyl sites on the chip surface were blocked by 1 M ethanolamine, pH 8. A flow chamber subjected to the immobilization protocol but without any immobilized protein was used as the control (blank) channel in each experiment.

A3G was sequentially diluted in running buffer (10 mm HEPES, 150 mm NaCl, 0.05% Tween-20, pH 7.4) and injected over into chambers at 30 μL/min. The binding experiments were monitored by the Biacore 3000 control software. Between experiments, the surfaces were strictly regenerated with 50 mM NaOH followed by an extensive wash procedure using running buffer. After normalization of the obtained data, the sensorgram from the control flow chamber of each injected concentration were subtracted. Binding curves were displayed, and the association (K_a), dissociation (K_d) rate constants, and affinities (K_D) were calculated using BIAevaluation 4.1 software.

SDS–PAGE and Western blot analysis

Proteins were analyzed by the 4-20% gradient SDS–PAGE (BIO-RAD) and stained for proteins with Coomassie blue. For Western blotting, protein samples were separated by SDS–PAGE and electrophoretically transferred to PVDF membranes at 50 V constant for 1 h using a Criterion™ Blotter (BIO-RAD). Membranes were incubated in PBS-T blocking buffer (PBS, 0.2% Tween-20, 5% BSA) at room temperature for 1 h. Blocked membranes were probed with either rabbit anti-A3G polyclonal antibody (Santa Cruz biotechnology) or mouse anti HIV-1 Vif monoclonal antibody (Santa Cruz Biotechnology) diluted 1:5000 in blocking buffer for 1 h at room temperature. After washing with PBST three times for 10 min each, the membranes were incubated with a horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (H + L) (1:5000, Pierce) or a rabbit anti-mouse IgG-HRP (1:5000, Santa Cruz Biotechnology) for 30 min and washed again. The presence of recombinant proteins was detected by NBT/BCIP Substrate (Sigma) according to manufacturer's instruction.

RESULTS AND DISCUSSION

Cloning and expression of Vif fusion proteins in *E. coli*

To generate the expression vector – pDEST527-His-Vif and pDest527-His-Vif$_{89\text{-}192}$, PCR-amplified full-length and truncated *vif* cDNAs were cloned downstream of 6×His tag by Gateway recombinant cloning procedure. In order to produce the His-Vif fusion proteins, the plasmids pDEST527-His-Vif and pDest527-His-Vif$_{89\text{-}192}$ were transformed and expressed in *E. coli* Rosetta 2 cells, respectively. Initial SDS-PAGE analysis showed that His-tagged Vif and

$Vif_{89\text{-}192}$ proteins were expressed as inclusion bodies. Several expression parameters, including post-induction temperature (18 or 37°C), time of induction (1-4 h for expression at 37°C, overnight for expression at 18°C), the concentration of IPTG (0.1–1mM), and different *E. coli* expression strains were evaluated in order to increase protein solubility. However, none of the conditions used showed a significant improvement in protein solubility (data not shown). The analyses of the insoluble fraction by SDS–PAGE showed His-tagged Vif is expressed with molecular weight of approximately 25 kDa corresponding to the fusion of 6×His tag (~1 kDa) and full-length HIV-1 Vif (~23 kDa); while the His-tagged $Vif_{89\text{-}192}$ is around 13 kDa, as shown in Figure 3.3. The identity of expressed proteins was confirmed by Western blot using the monoclonal anti-6×His and anti-Vif antibodies.

Purification and refolding of His-Vif fusion proteins

Inclusion bodies containing His-Vif and His-$Vif_{89\text{-}192}$ in the cell lysate were pelleted by centrifugation and washed with urea and Triton X-100 to remove *E. coli* membrane and cell wall material. Urea concentration of 2 M was selected to be included in the wash buffer because higher urea concentration may result in partial solubilization of the recombinant proteins. We also included 2% Triton X-100 in the wash buffer to help extract lipid and membrane-associated proteins. A final wash step with the detergent free buffer ensures removal of excessive detergent from the inclusion bodies. 6 M GnCl was used to solubilize His-Vif in this study, because it appeared to be more efficient in solubilizing His-Vif fusion proteins than 8 M urea (data not shown). An extensive purification was achieved using nickel affinity columns as shown from the chromatogram and SDS-PAGE analysis (Figure 3.2). Typically, approximately 10 mg of denatured His-Vif or His-$Vif_{89\text{-}192}$ was produced from one gram bacterial cell pellets. The SDS–

PAGE insert in Figure 3.2 shows a pool of the peak elutes of the His-Vif stained with Coomassie blue. The purity achieved from this purification step is estimated to be ~90%, and the eluted proteins were confirmed by Western blot using the anti-6His and anti-Vif antibodies.

Since the Vif fusion proteins were purified from inclusion bodies as denatured proteins in the presence of 6 M GnCl, they have to be refolded for further structural or functional studies. The refolding of recombinant Vif proteins was accomplished by a stepwise dialysis with decreasing GnCl concentration from 6 M to 0 M as described in Experimental methods. The overall yield of each refolded Vif fusion protein was approximately 10 mg/L of bacterial cultures.

Our results also revealed that both refolded full-length Vif and truncated $Vif_{89\text{-}192}$ proteins expressed from *E. coli* had a strong tendency to form multimers. Specifically, refolded Vif fusion proteins were boiled in Laemmli sample buffers containing 2% SDS and 5% β-mercaptoethanol (BME) for 5 min and loaded onto a 4-20% Tris/glycine SDS-PAGE for electrophoresis. Under this denaturing condition, His-Vif proteins migrated in the gel as monomers (25 kDa) and dimers (50 kDa). The majority of His-Vif was in the form of monomer as shown in Figure 3.4. In contrast, under the same denaturing condition, the SDS-PAGE analysis showed that His-$Vif_{89\text{-}192}$ existed as monomers (13 kDa), dimers (26 kDa), trimers (39 kDa), tetramers (52 kDa), pentamers (65 kDa), and even higher-order multimers up to 10-mers or higher, as shown in Figure 3.5. a. We also tested the stability of His-$Vif_{89\text{-}192}$ multimers by incubating the protein with various additives including 5% SDS, 500 mM DTT, 2% Triton X-100, Brij 35, Zwittergent, 6 M urea, 0.1 M EDTA, 5% DMSO, and 0.1 mg/mL of DNase and RNase at room temperature for 30 min respectively, and then boiled in Laemmli sample buffer for 5 min before subjecting to SDS-PAGE. Our results showed that the His-$Vif_{89\text{-}192}$ multimers were very stable and could

withstand harsh conditions as detergent, reducing agent, denaturing agent and heat (data not shown). Of note, this multimerization is also DNA- and RNA-independent. To further investigate the multimerization of truncated Vif_{89-192}, we expressed and purified GST- and MBP-tagged Vif_{89-192} from *E. coli*. The SDS-PAGE analysis showed that these two recombinant Vif proteins could also form higher-order multimers up to 10-mers and exhibited a very similar multimerization pattern as His-tagged Vif_{89-192} (data not shown). Besides the bacterial expressed Vif_{89-192} fusion proteins, we also expressed His-Vif_{89-192} in baculovirus system using insect Sf9 cells. Although the protein was partially soluble and failed to purify, Western blot of cell lysate showed the Vif_{89-192} expressed from insect cells could form multimers as well (data not shown).

It has been reported by several independent studies that Vif can form multimers and the multimerization motif is located in the C-terminal domain [15-17]. Moreover, the Vif multimerization is crucial for its function and HIV-1 infectivity [18-22]. Besides Vif, several other HIV-1 accessory proteins such as Rev, Nef, Vpu, Vpr and Tat, can also form dimers and/or oligomers. For example, the HIV-1 Rev that functions protein functions in the HIV-1 mRNA nuclear transport can form multimers of eight molecules or more through a combination of cooperative protein-protein and protein-RNA interactions [23]. In addition, it has been suggested that dimerization is required for HIV-1 Nef activities in CD4 down-regulation [24]. Also, another HIV-1 accessory protein - Vpu can form high-molecular-weight complexes consisting of homo-oligomers [25], which is essential for tetherin binding [26]. However, Vpr can form dimers and oligomers and the oligomerization is essential for its incorporation into virus particles [27]. Tat forms a metal-linked dimer with metal ions bridging cysteine-rich regions from each monomer and it is important for Tat functions [28, 29]. Other than HIV-1, other viral proteins can also form multimers, such as Ebola virus Vp30 [30], Influenza virus NS1 protein [31],

measles virus attachment protein H [32], Rift valley fever virus L protein [33], hepatitis delta antigen of hepatitis delta virus [34], and so on. If the multimerization of non-structural viral proteins is not ubiquitous, at least it is not unusual. In this study, the recognizable difference of multimerization pattern between full-length Vif and truncated Vif_{89-192} remains a question. The mechanism that causes the phenomenon and the question whether there is a physiological meaning underlying the difference require further investigations.

However, our results suggest that not only the Vif C-terminal PPLP motif but also the residues from N-terminal region were involved in Vif multimerization, possibly through a negative cooperative binding manner. The binding of Vif N-terminus to its partners may somewhat promote the Vif C-terminal self-association activity to form higher-order multimers. This A3G-induced formation of Vif higher-order multimers, in turn, enhance the binding capability for more A3G molecules, which help explain how Vif can effectively mediate the formation and degradation of high molecular mass (HMM) complex of A3G [35]. Furthermore, we have identified a small molecule, VMI-1, that can antagonize Vif_{89-192} multimerization. The addition of this compound at the concentration of 1 mM could effectively inhibit the formation of Vif_{89-192} multimers as shown in Figure 3.5. b and c, suggesting that the Vif-Vif self-association is unlikely to be covalent.

Expression and purification of recombinant A3G

Initially, we attempted to express recombinant full-length A3G in *E. coli*. However, in bacterial systems the expression level of the recombinant protein was too low to be purified. Therefore, we resort to the Bac-to-Bac baculovirus expression system. To facilitate purification, two fusion tags, 6×His and MBP are placed at the N-terminus of A3G with the TEV protease

recognition sequence in between. The Optimized expression condition was tested by harvesting the cells at 48 and 72 h post-infection, and by infecting the cells at different MOI. The highest protein production was found when cells were harvested at 72 h post-infection and the expression level showed no significant increase with the increasing multiplicity of infection (MOI) from 0.5 to 5. Therefore, the selected optimal condition for large-scale A3G expression and subsequent purification was to harvest cells 72 h post-infection with MOI around 0.5. After purification using two purification columns and TEV cleavage, the A3G exhibits a major single band (>96% purity) with a molecular weight of 45 kDa on SDS-PAGE, and its identity was confirmed by Western blot using anti-A3G antibody as shown in Figure 3.6. Overall, a single His-tag affinity column was unlikely to purify the A3G from crude cell lysate to an anticipated purity. Thus an additional MBP affinity purification step removed most of the unwanted contamination proteins, therefore provided a preferred purity for further studies. Typically after the final gel filtration purification step, the total yield of high purity A3G was 1 mg per liter of insect cell culture. It is worth noting that gel filtration chromatography isolated A3G in both *monomeric* and dimeric forms with the elution volume of 8.7 mL and 12.1 mL, respectively, on a Superdex 75 10/300 column as shown in Figure 3.6.

It has been observed that A3G can exist as an inactive, high molecular mass (HMM) complex that can be transformed *in vitro* into an active, low molecular mass (LMM) variant [36]. In addition, A3G contains two cytidine deaminase domains (CDAs), of which only the C-terminal domain is known to have cytidine deamination activity. The oligomeric state of A3G is therefore proposed to support A3G's uneven deaminase activity. By modeling the crystal structure of the related tetrameric APOBEC2 (A2) protein, it has been proposed that A3G can form dimers [37]. Additional evidence derived from atomic force microscopy, mutagenesis

analysis and SAXS also suggest the existence of A3G dimers in solution [38-43]. Subsequent mutagenesis analysis suggested that the dimerization of A3G may be nucleic acid-dependent and essential for its deaminase activity on ssDNA [44]. Although the detailed interface of A3G dimerization has not been identified, it is believed that A3G N-terminal domains symmetrically interact to each other via a head-to-head conformation. In this study, we reported, for the first time, the isolation of stable A3G in the dimeric conformation with a manageable amount. The purified A3G dimers can be used not only to elucidate the interface of A3G self-association, but also to investigate its physiological significance.

The dimerization and formation of higher-order oligomers by protein self-association are very common. Many important proteins, such as receptors (G-protein-coupled receptors [45], tyrosine kinase receptors [46]), ion channels [47], and enzymes can form homodimers. It has been reported that ~70% of enzymes found in the Brenda database (http://www.brenda-enzymes.org/) can self-associate to form dimers or higher-order oligomers [48]. Protein dimerization and oligomerization can provide various structural and functional advantages, such as increase stability, fine-tune the accessibility and specificity of active sites, and increase protein complexity [48, 49]. Like many other RNA and ssDNA binding proteins, A3G is likely to be in the oligomerization state to facilitate the nucleic acids binding. From our gel filtration chromatography result, the A3G is possibly in a dynamic equilibrium between monomers and dimers, while the predominant conformation is A3G monomers. This leverage may shift toward dimerization upon binding substrate RNA and ssDNA.

AlphaScreen studies

To evaluate the interaction between Vif and A3G, we performed the AlphaScreen binding assay. In an established AlphaScreen experiment, the 680 nm laser excitation of donor

beads produces singlet oxygen molecules. If the acceptor beads are within a close proximity of approximately 200 nm with donor beads, the energy from singlet oxygen molecules transfer to acceptor beads to generate a light signal of 620 nm. Donor and acceptor beads are engineered with functional groups that can be bioconjugated with specific analytes. Here, we have configured an AlphaScreen assay by using His-Vif coated donor beads and A3G-coated acceptor beads to measure the binding of these two proteins, as shown in Figure 3.7. In our assay, different amount of refolded His-Vif (10-150 ng) was first incubated with Ni^{2+}-chelate donor beads, respectively, to allow the association of proteins to the beads. Afterwards, titration of A3G-conjugated acceptor beads at the equal concentration with donor beads initiated the binding of two proteins to occur and the light signals were recorded. Our results showed that when 50 ng of His-Vif was present in the donor beads solution, the maximum signal production (300,000 AlphaScreen response units) was observed, as shown in Figure 3.8. When a higher concentration of His-Vif (75 - 150 ng) was present, a significant decrease in the signal was observed due to the so-called "hook effect" [50], whereby His-Vif concentration surpasses the binding capacity of the donor beads, and the donor beads became saturated with conjugated proteins. The excessive unconjugated His-Vif proteins present in the reaction compete with conjugated proteins for binding with A3G in a competitive inhibition mechanism resulting in the decrease in signal. In contrast, when less than 50 ng His-Vif were present in the reaction (10 and 25 ng), the signals increased in proportion with His-Vif concentration. Our results showed that when less His-Vif was conjugated with the donor beads, less donor–acceptor beads were brought to close proximity and signals decreased accordingly. Overall, AlphaScreen results provided qualitative evidence for the direct binding of full-length Vif with A3G *in vitro*.

His pull-down assay

We used protein affinity His pull-down assays as an alternative approach to determine whether His-Vif and A3G interact directly. *E. coli*-expressed His-Vif fusion protein was purified and bound to Ni Sepharose resin. The Vif conjugated beads were collected in an empty column and washed extensively. Purified A3G was then added to the column to allow the binding to take place. After incubation, the Ni resin was extensively washed to remove any unbound A3G. The A3G flow-through fraction and beads were boiled in Laemmli loading buffer, and bound proteins were detected by SDS-PAGE. The results showed additional evidence that Vif bound to A3G directly as indicated in Figure 3.9, and provided additional evidence for a direct interaction between full-length Vif and A3G.

Biacore surface plasmon resonance

To further analyze the binding properties of the Vif and A3G proteins and to establish the affinity constant of the binding, the Biacore 3000-Surface Plasmon Resonance (SPR) analysis was performed. The Biacore 3000 is an optical biosensor that uses SPR for real-time determination of the kinetics and affinities of protein-protein interactions without the use of labels. In our experiment, the His-Vif and His-$Vif_{89\text{-}192}$ were chosen to be immobilized on the surface of CM5 sensor chip flow chambers via amine coupling, respectively. Subsequently, various concentrations of A3G monomers and dimers were injected into the Vif-immobilized chambers to determine kinetics constants. The K_D (affinity) was calculated from the determined K_a and K_d values. For the interaction analyses, both A3G monomers and dimers were analyzed respectively. In our experiment, A3G monomer (2.5 μM, 5 μM, 10 μM, 20 μM, and 40 μM) dissolved in running buffer were injected into the full-length or truncated Vif immobilized

channels using a flow rate of 15 μL/min for 3 minutes to monitor association of the Vif ligand and A3G analyte. After the injection step, running buffer was injected over the surface to monitor the dissociation. The K_D for the interactions (summarized in Table 3.2) between A3G monomer and full-length Vif was 200 nM (Figure 3.10 a), while the A3G monomer – $Vif_{89\text{-}192}$ binding had a K_D of 200 mM (Figure 3.10 c). Therefore, our Biacore analysis showed the *in vitro* A3G-Vif binding is of high affinity, and A3G directly interacts with N-terminal region of Vif as indicated by the low affinity between $Vif_{89\text{-}192}$ and A3G. In addition, we also measured the binding kinetics between A3G dimer and Vif proteins. Surprisingly, the A3G dimers exhibited an affinity to the full-length Vif almost 1000 fold greater than A3G monomers, with a calculated k_D of 0.2 nM, (Figure 3.10 b). In contrast, the k_D of A3G dimer – $Vif_{89\text{-}192}$ binding was 240 mM that is very similar to the Kd of A3G monomer – $Vif_{89\text{-}192}$ binding (200 mM). Again, our results showed that neither A3G dimer nor monomer bound to Vif C-terminal regions, which suggested that Vif N-terminal region is responsible for the strong interaction between A3G dimer and full-length Vif. However, in order to determine meaningful kinetic constants in a direct protein–protein interaction Biacore experiment, the analyte protein should be a monomer, or at least behave like one, so that the data can be modeled with a simple bimolecular reaction mechanism [51, 52]. Thus, in our Biacore experiments, when we calculate the kinetics constants between A3G dimers and Vif proteins, each A3G dimer was considered as one functional unit and the molar concentration was calculated accordingly, i.e. 1 mM A3G dimer equals to 2 mM A3G monomer and the molecular weight of A3G dimer is 90 kDa instead of 45 kDa of A3G monomer. Overall, our Biacore results indicated that both A3G monomer and dimer interact directly with full-length Vif by binding to the N-terminal region of Vif, and A3G dimers exhibited a 1000 higher Vif-binding affinity than A3G monomers.

Because of the lack of pure and stable A3G dimers, it has been questionable whether the dimerization of A3G affects the Vif binding over recent studies. Lavens *et al.,* found that mutations in the proposed A3G N-terminal dimerization region strongly reduce A3G dimerization and binding with Vif as tested using mammalian protein–protein interaction trap two-hybrid technique [53]. However, it was also reported the degradation of A3G by Vif was not affected by mutations in the A3G N-terminal dimerization region [37, 54]. To date, two distinct models of A3G-Vif binding had been proposed. The "A3G monomer-Vif" model suggests that Vif binds to the A3G monomer at the binding site used for A3G dimerization and the binding of Vif inhibits the dimerization of A3G. The second model depicts that Vif binds to the A3G dimer and the binding requires A3G dimerization [53]. Different from above mentioned models, our Biacore results an A3G-Vif binding mechanism that Vif could bind to both A3G monomer and dimer, but Vif preferably interacts with A3G dimers with an affinity of 1000 times greater than A3G monomer. The correctness of this model still needs to be validated by studies using other biophysical analyses, especially certain competitive binding analyses. Another question that needs to be addressed is the stoichiometry of A3G-Vif binding. Based on our results that Vif prefers A3G dimers to monomers while Vif itself exists as multimers in solution (see chapter 2 for details), the binding of A3G with Vif may promote the self-association of Vif C-terminal multimerization motif. However it has been suggested that Vif C-terminal region is very likely transformed from unstructured to structured upon binding with other biological partners, such as E3 ubiquitin ligase complex [11]. In addition, the only available structure of Vif is a short fragment of α-helix in Vif C-terminal domain when complexed with E3 ubiquitin ligase complex [10]. All these data strongly suggest that Vif is likely to be structured upon interaction with other proteins [12]. However, little is known about the Vif binding with E3 ubiquitin ligase complex.

We still do not know whether the binding would diminish Vif self-association, however it is very possible. In order to study Vif-A3G interaction, E3 ubiquitin ligase complex and/or other Vif C-terminal binding partners must be taken into the experimental system. The functions and interactions that Vif confers in the HIV-1 biology are very likely more complex and multifaceted than what we have known so far. Presently, a high-resolution structure of Vif-A3G interface is of primary importance in defining the interaction between these 2 proteins and will cast light on the future drug design aimed at blocking their interactions.

CONCLUSION

HIV-1 Vif directly interacts with the host antiviral factor - APOBEC3G and leads it for degradation through the proteasome pathway. The A3G degradation requires the binding of the cellular E3 ubiquitin ligase complex with Vif, which largely consists of cullin5, elongin B, and elongin C. The elongin C directly binds to Vif C-terminal SOCS BC-Box motif [55], while Cullin5 binds to Vif HCCH motif [14, 56] flanking BC-Box motif. On Vif-mediated A3G degradation, Vif therefore is serving as an adaptor molecule to recruit the E3 ubiquitin ligase complex to its substrate – A3G, the E3 ubiquitin ligase complex induces the polyubiquitination of A3G and directs it to the 26S proteasome for degradation. As a result, this function of HIV-1 Vif makes it a critical factor in HIV-1 infectivity and virus spread, and provides a potential pharmacological target.

In this study, we analyzed the interaction between full-length HIV-1 Vif protein and human A3G using biophysical and biochemical methodologies and provided evidence for the direct binding of Vif with A3G under physiological conditions. The binding kinetics of the interaction was also measured using surface plasmon resonance. However, the difficulty of

expression and purification of full-length Vif in prokaryotic and eukaryotic systems with a useful quantity for structural and functional studies has greatly encumbered the understanding of Vif functions. In our study, we purified and refolded full-length Vif with a satisfactory quantity and purity. Our results also revealed that both refolded full-length Vif and truncated Vif_{89-192} proteins expressed from *E.coli* have a strong tendency to form multimers. Under denaturing conditions, Vif migrated in the gel as monomers (25 kDa) and dimers (50 kDa), while the majority of His-Vif was in the form of monomers. In contrast, Vif_{89-192} was not only able to form monomers and dimers, but also trimers (39 kDa), tetramers (52 kDa), pentamers (65 kDa), and even higher-order multimers. We also tested the stability of His-Vif_{89-192} multimers and our results showed that the His-Vif_{89-192} multimers were very stable and can withstand such harsh conditions as detergent, reducing agent, denaturing agent and heat. Further expressed and purified GST- and MBP-tagged Vif_{89-192} from *E. coli* showed a similar stable multimerization pattern. In addition, His-Vif_{89-192} protein expressed from baculovirus could form multimers as well. Based on our results, we speculate that the existence of Vif N-terminal sequence could inhibit Vif_{89-192} self-association. The mechanism that causes the differential multimerization is worth further investigation.

During our gel filtration chromatography purification of A3G, we consistently observed the elution of dimeric A3G with the MW of 90 kDa, which is twice the relative molecular weight of monomeric A3G. No higher-ordered species of A3G was observed in our study. We have, of note, for the first time purified stable A3G dimers that provides direct evidence that A3G could form dimers [37].

Importantly, we have measured the binding kinetics of A3G monomers and dimers with full-length and truncated Vif, respectively. Our results showed that full-length Vif preferably

interacts with A3G dimers (K_D=0.2 nM) with an affinity of 1000 times greater than with A3G monomers (Kd=200 nM), whereas both A3G monomers and dimers did not show any significant affinity to truncated $Vif_{89\text{-}192}$ (KD=200 mM). Therefore, we proposed that although Vif can possibly bind to both A3G monomer and dimer, dimeric A3G has the preferred conformation for Vif binding. Ultimately, future attempts for the high-resolution atom-level structure of Vif-A3G dimer complex may help efforts aimed at therapeutic intervention of the interaction between the HIV-1 Vif protein and A3G.

Figure 3.1. The functional domains in HIV-1 Vif. Known HIV-1 Vif functional domains are aligned with the PONDR VL-XT prediction of Vif order/disorder regions [57-59]. Interaction partners are described along the top of the Figure. The gray line in the PONDR plot represents the PONDR score of 0.5: regions with scores below 0.5 are likely to be structured, while a score above 0.5 indicates unstructured regions. The amino acid residue numbers are at the bottom of the plot, which is in accordance with the relative position of each major Vif function domain. This figure and the information therein have been adapted from [13, 60]

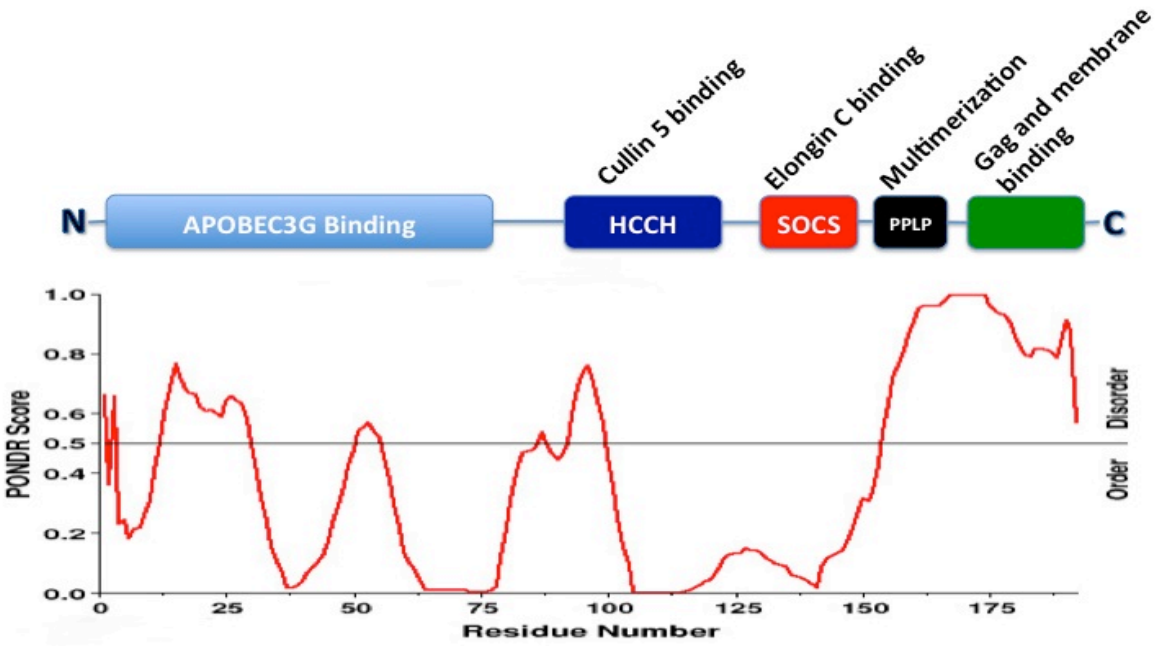


Table 3.1. The sequences of primers used in this study. The DNA oligonucleotides were used to generate the expression constructs of full-length Vif, truncated Vif_{89-192} and full-length A3G. Vif and Vif_{89-192} primers contain Gateway® cloning specific recombinant sequence. A3G primers contain EcoRI and XbaI restriction sites respectively, and the TEV protease recognition sequence.

DNA oligonucleotides	Sequence
Vif for	GAAAACCTGTACTTCCAAGGCGGGTCAGGTATGGAAAA CAGATGGCAGGTGATGA
Vif rev	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGTGTCCA TTCATTGTATGGC
Vif_{89-192} for	GAAAACCTGTACTTCCAAGGCGGGTCAGGTATGTGGAG GAAAAAGAGATATAGCA
Vif_{89-192} rev	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGTGTCCA TTCATTGTATGGC
A3G for	GAAAACCTGTACTTCCAAGGCGGGTCAGGTAAGCCTCAC TTCAGAAACAC
A3G rev	GGGGACCACTTTGTACAAGAAAGCTGGGTTCAGTTTTCC TGATTCTGGAGAATG

Figure 3.2. Flow chart of purification of recombinant A3G from insect Sf9 cells.

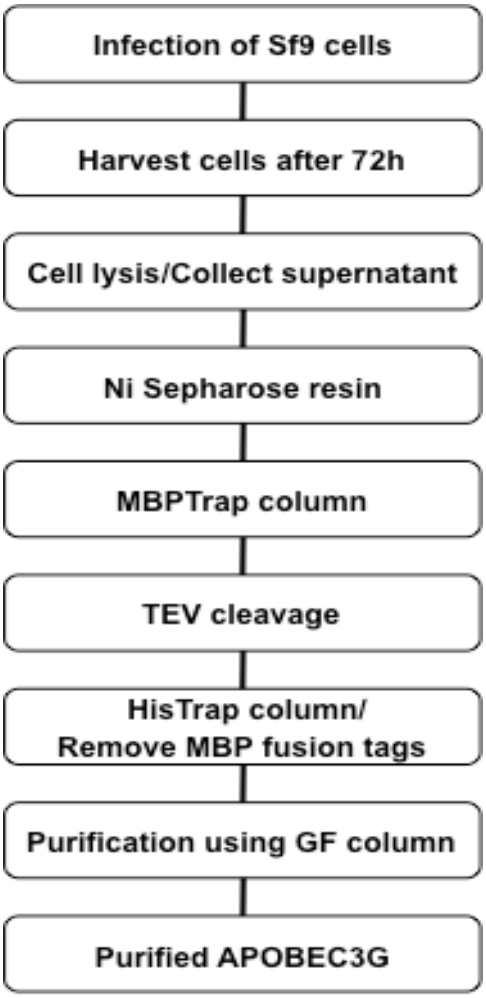


Figure 3.3. Bacterial expression and purification of recombinant full-length Vif and truncated $Vif_{89\text{-}192}$.

a, SDS–PAGE analysis of expression of two recombinant Vif proteins in *E.coli*, uninduced cells (1) and induced cells (2) for truncated $Vif_{89\text{-}192}$ expression, recombinant truncated $Vif_{89\text{-}192}$ was over-expressed with the MW of ~13 kDa. Uninduced cells (3) and induced cells (4) for full-length Vif expression, recombinant full-length Vif was over-expressed with the MW of ~25 kDa.

b, Western blot of SDS-PAGE in panel a using anti-Vif antibody.

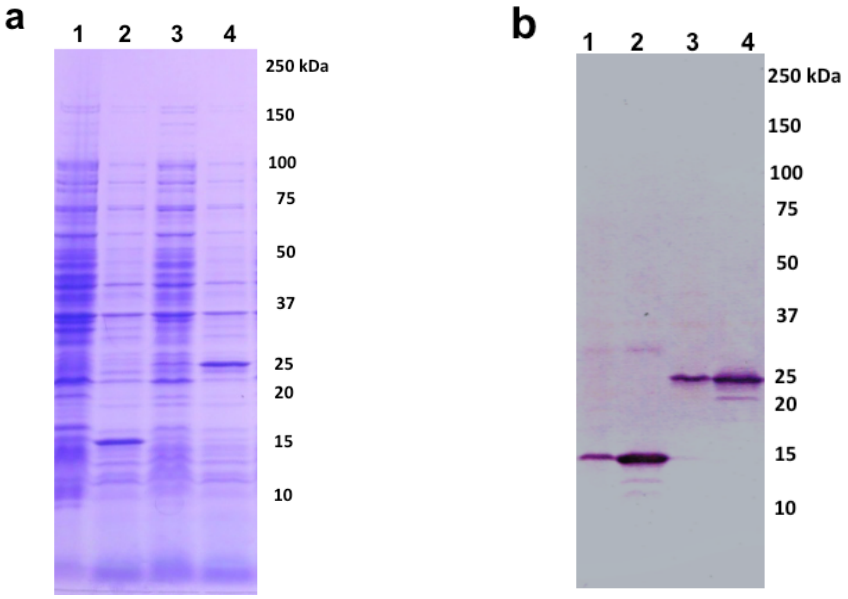


Figure 3.4. Purification and refolding of recombinant full-length Vif.

a, chromatogram of HisTrap column purification of His-tagged full-length Vif in denaturing condition (blue line). The His-tagged Vif was eluted by an imidazole gradient (20 – 500 mM, green line); insert shows the SDS-PAGE analysis of elute fractions containing purified Vif.

b, SDS-PAGE analysis of refolded Vif showing Vif monomer (25 kDa) and Vif dimer (50 kDa) as indicated by arrows.

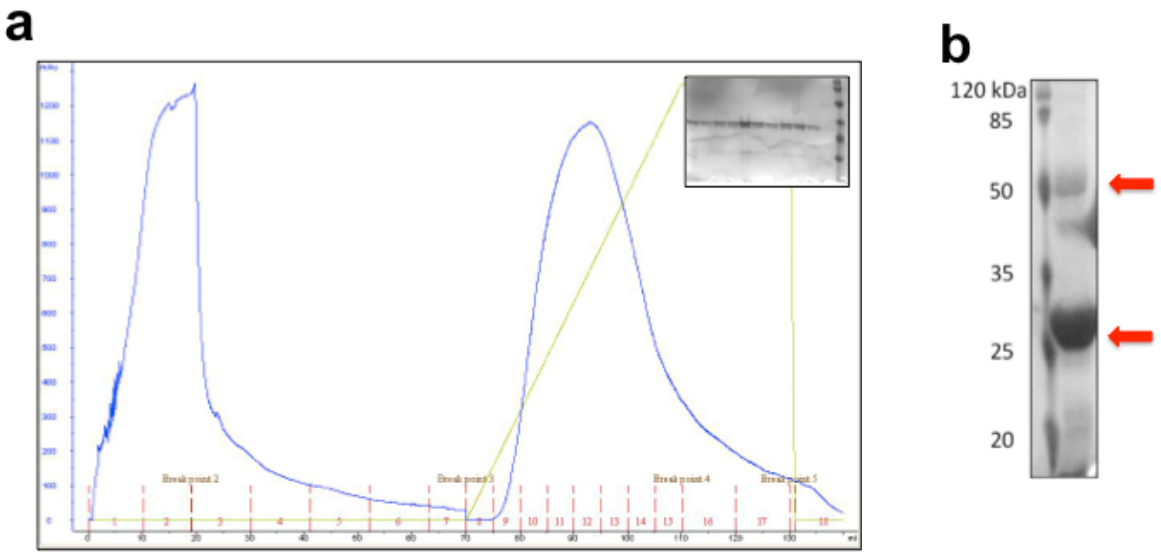


Figure 3.5. Multimerization of truncated Vif_{89-192}.

a, Left panel: SDS-PAGE analysis of refolded Vif_{89-192} showing monomers (~13 kDa), dimers (~26 kDa), trimers(~35 kDa), tetramers (~52 kDa), pentamers (~65 kDa) and higher-order multimers as indicated by arrows. Right panel: Western blot of Vif_{89-192} multimers using anti-Vif antibody.

b, Left panel: SDS-PAGE analysis of refolded Vif_{89-192} after VMI-1 treatment （lane 1）and before treatment (lane 2), showing that VMI-1 can effectively inhibit the multimerization of Vif_{89-192}. Right panel: Western blot of SDS-PAGE on left panel using anti-Vif antibody.

c, chromatogram of refolded Vif_{89-192} (after VMI-1 treatment) using a Superdex 75 10/300 gel filtration column in native condition (blue line). The Vif_{89-192} was eluted in a broad peak containing monomers and dimers.

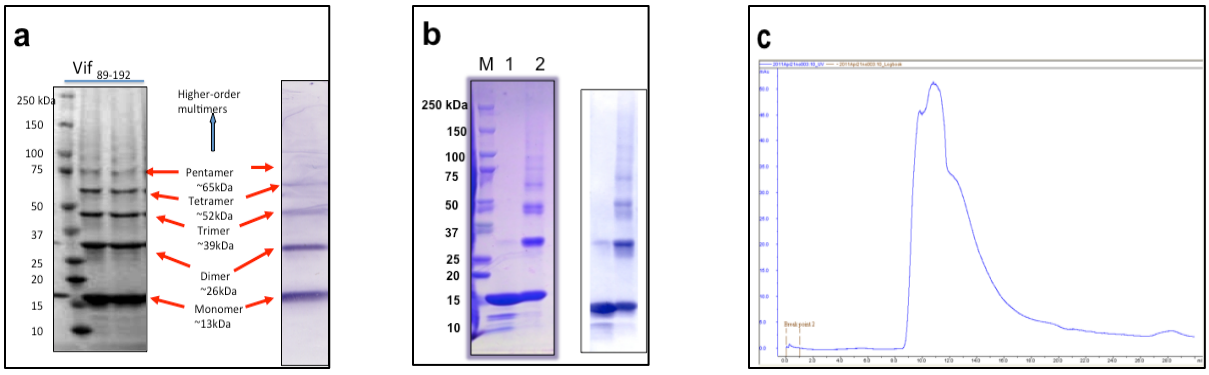
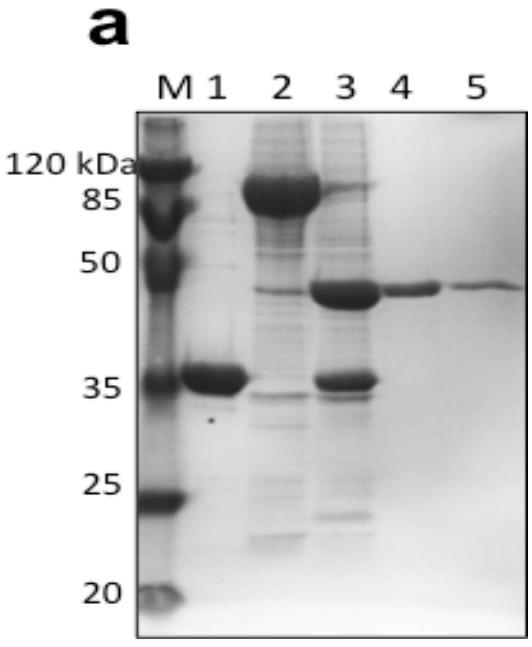


Figure 3.6. Purification of A3G from insect Sf9 cells.

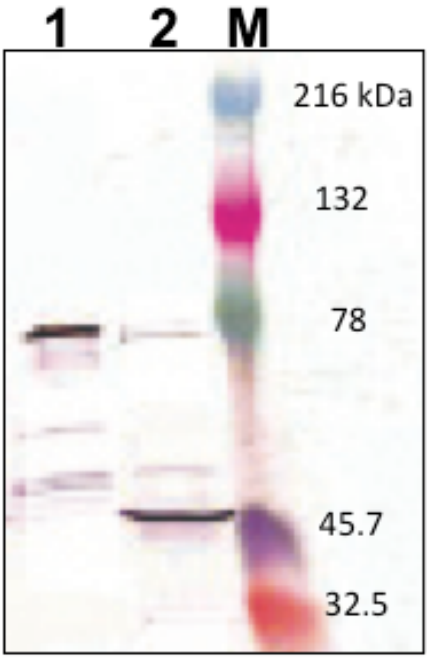
a, purification of A3G by SDS-PAGE analysis. M: molecular weight markers. 1: TEV protease. 2: recombinant tagged A3G purified with HisTrap and MBPTrap columns. 3: purified recombinant tagged A3G incubated with TEV protease overnight. 4-5: the flow-through fractions of 2nd round HisTrap column purification, showed the purified A3G about 45 kDa.

b, Western blot of recombinant A3G before fusion tag removal (lane 1) and after fusion tag removal (lane 2) using anti-A3G antibody.

c, chromatogram of A3G purification using a Superdex 75 10/300 gel filtration column (blue line). The A3G protein was eluted at 8.7 mL and 12.1 mL respectively, corresponding to the MW of ~90 kDa and ~45 kDa. Insert, SDS-PAGE analysis of fractions in each elution peak shows identical protein MW.



b


1
2
M
216 kDa
132
78
45.7
32.5

c

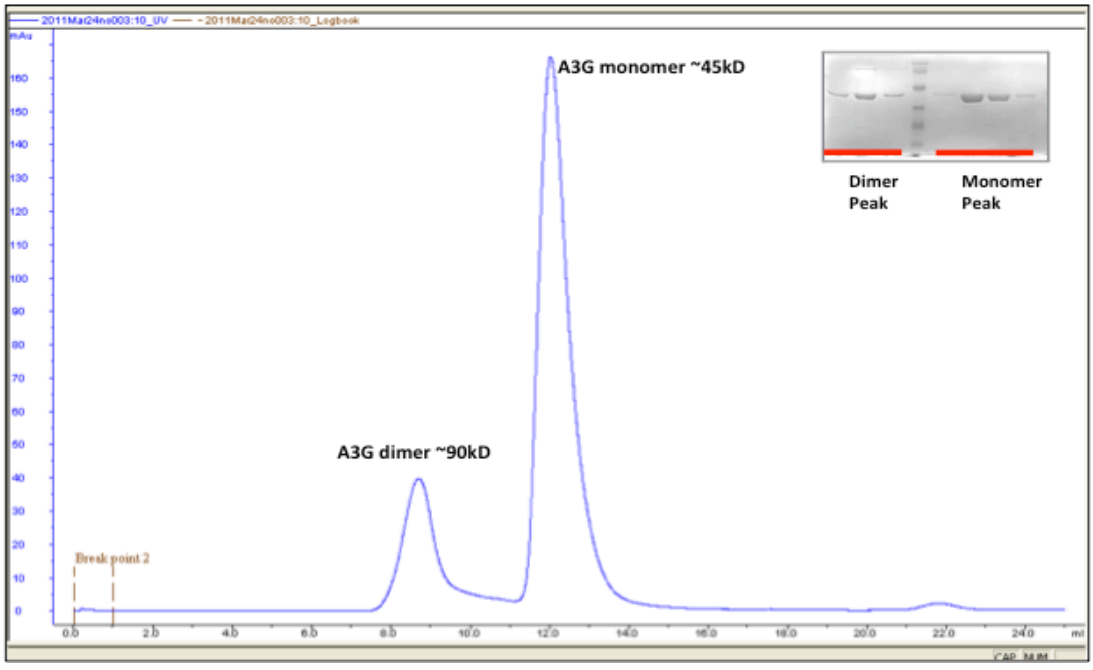

A3G monomer ~45kD
A3G dimer ~90kD
Dimer Peak
Monomer Peak
Break point 2

Figure 3.7. Schematic representation of the AlphaScreen assay as used in this study. His-tagged Vif was bound to Nickel Chelate Donor beads, while A3G conjugated to Acceptor beads covalently. When the Vif binds to the A3G, the two beads are brought into close proximity. Excitation of the donor bead at 680 nm causes the generation of singlet oxygen that can diffuse up to 200 nm in solution before reacting with acceptor beads. If the acceptor beads are within this range, as would be in the case of specific binding, it will produce an emission at 620 nm, which can be detected by EnVision reader.

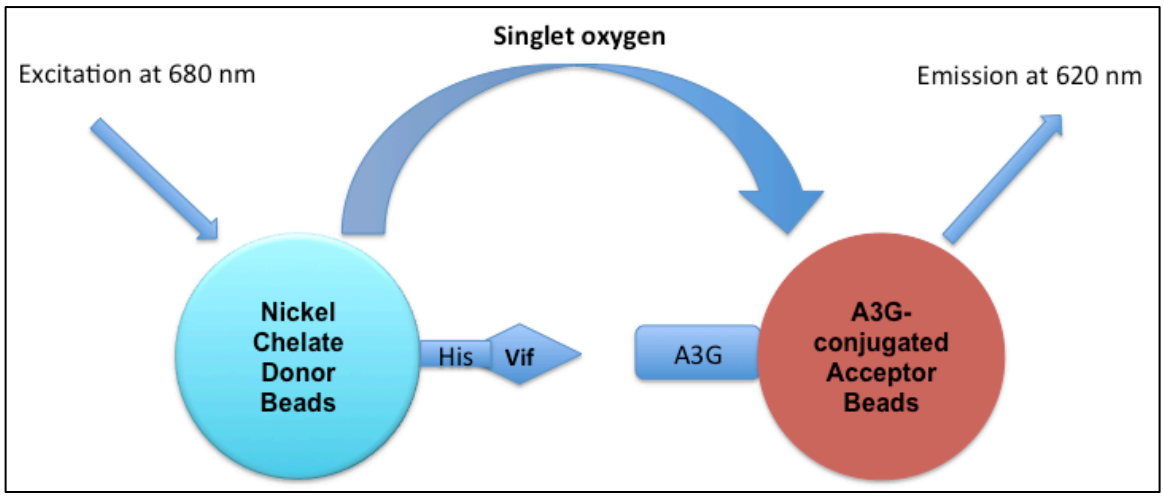
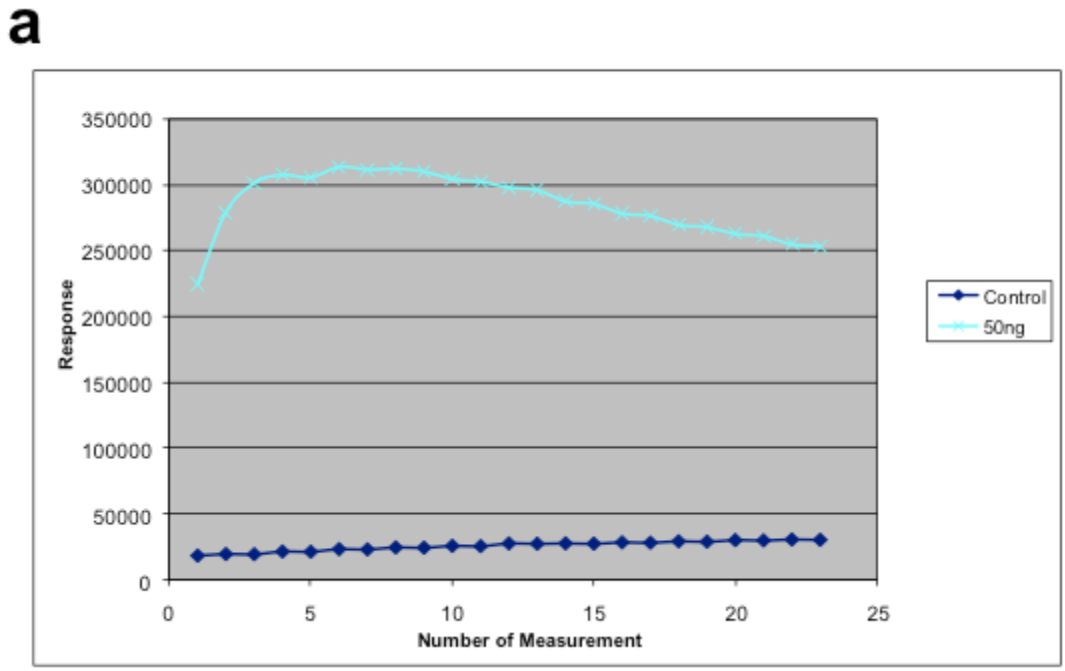


Figure 3.8. AlphaScreen analysis of Vif-A3G binding.

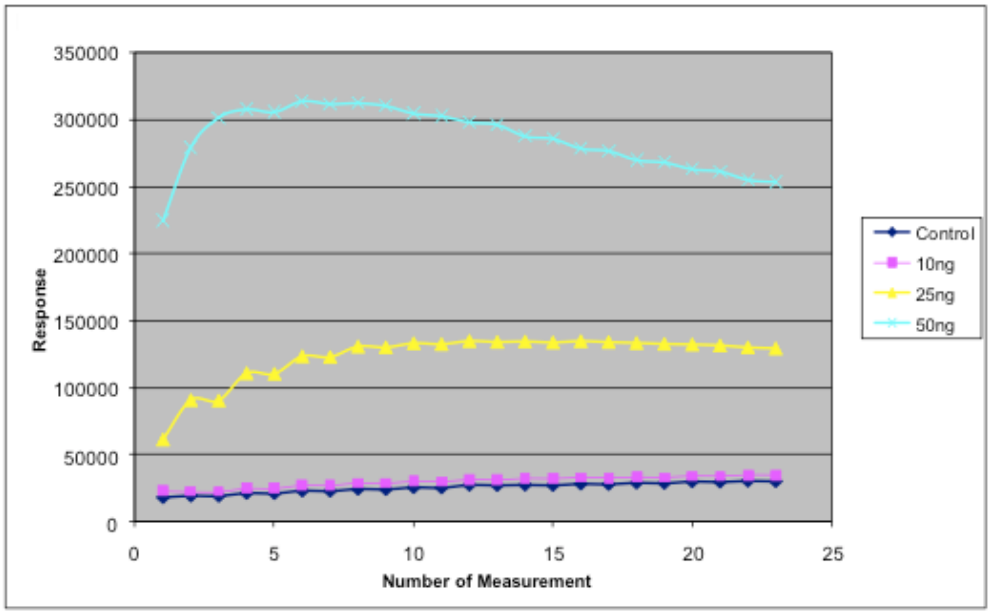
a, AlphaScreen results showing the interaction between Vif and A3G. The maximum signal production appeared when 50ng His-Vif was added (cyan line); the strong signal indicates the binding of APOBEC3G with Vif.

b, when less than 50ng of His-Vif was added, signal intensities decreased accordingly, indicating a limited association of Acceptor and Donor beads took place.

c, when more than 50ng of His-Vif was added, decreasing signal intensities were observed due to the fact that excessive unconjugated His-Vif in solution competes with binding sites on Acceptor beads; therefore less signal is generated.



b


350000
300000
250000
200000
150000
100000
50000
0
Response
0
5
10
15
20
25
Number of Measurement
Control
10ng
25ng
50ng

c

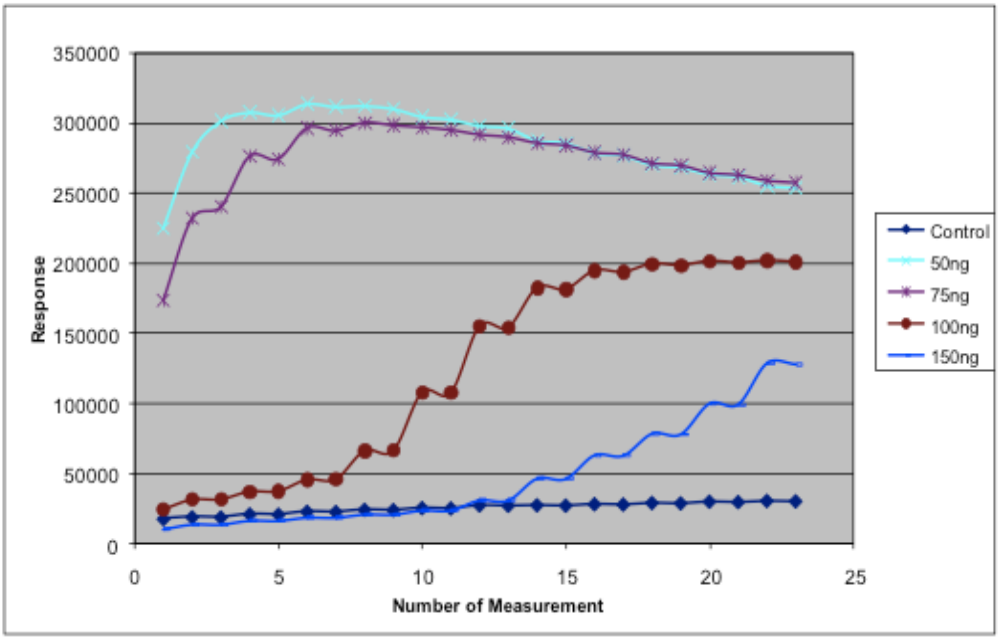

350000
300000
250000
200000
150000
100000
50000
0
Response
0
5
10
15
20
25
Number of Measurement
Control
50ng
75ng
100ng
150ng

Figure 3.9. SDS-PAGE analysis of His pull-down assay.

1: Purified full-length His-tagged Vif.

2: Purified A3G.

3: Vif bound Ni-NTA resin after incubation with A3G, showing the presence of both Vif and A3G proteins.

4: Flow-through after A3G incubation showing excessive A3G.

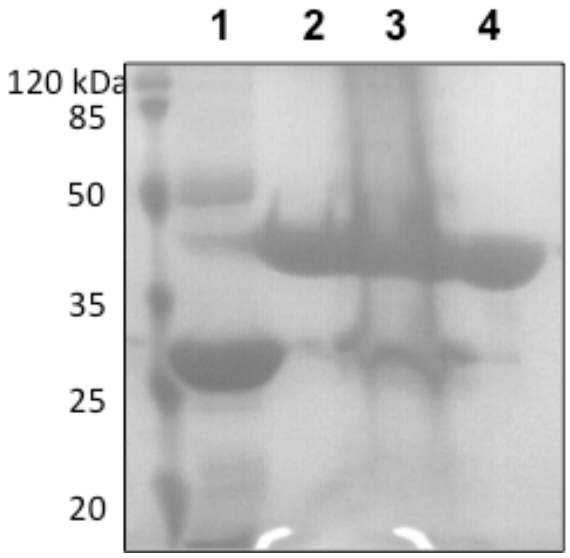


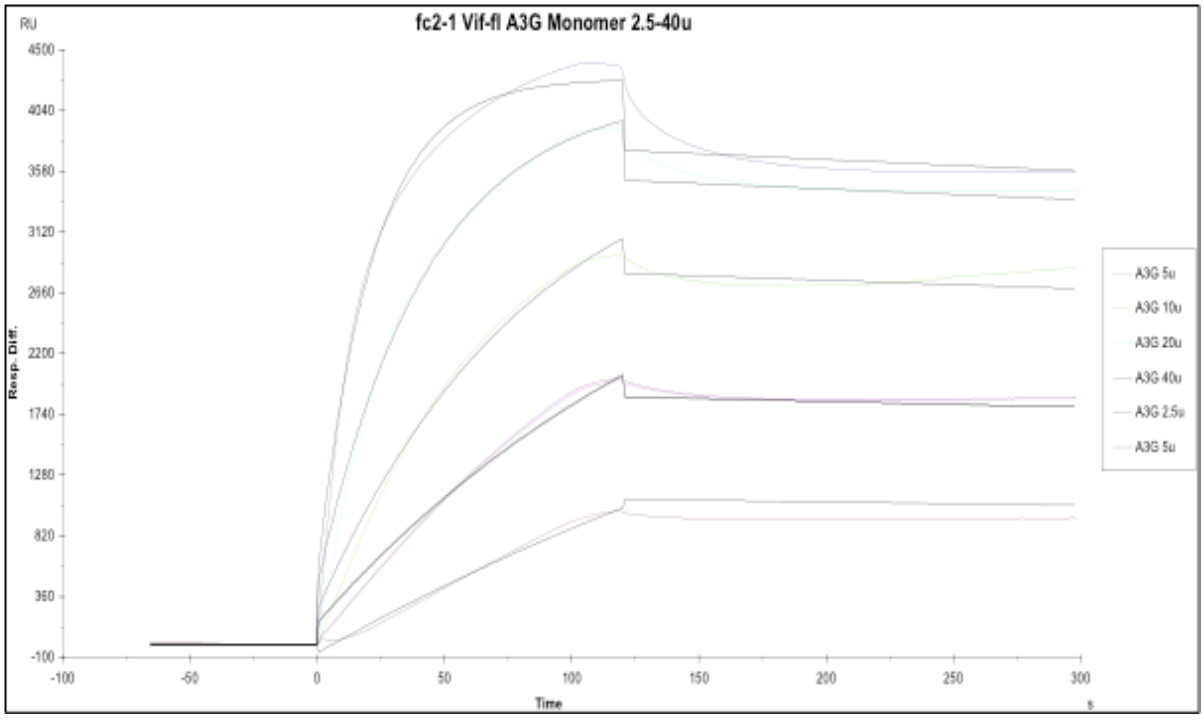
Figure 3.10. Biacore surface plasmon resonance studies on Vif-A3G binding.

a, sensorgram used for kinetic analysis of interaction between A3G monomer at different concentration (2.5 - 40 μM) and immobilized full-length Vif.

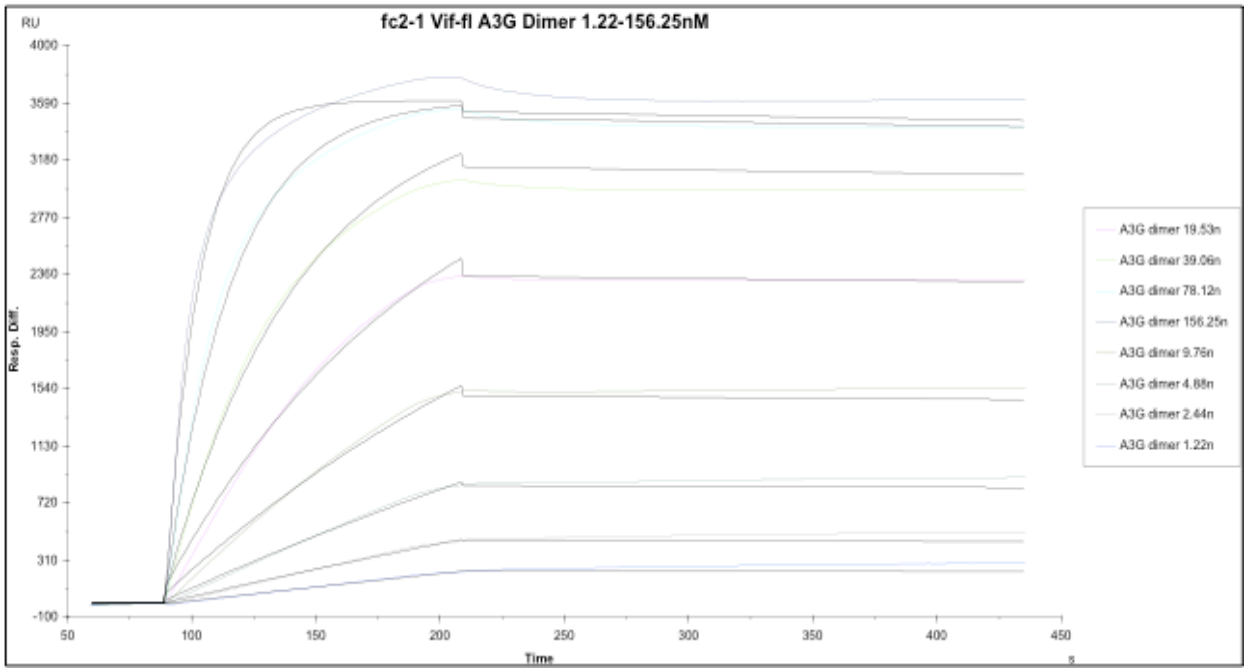
b, sensorgram used for kinetic analysis of interaction between A3G dimer at different concentration (1.22 – 156.25 nM) and immobilized full-length Vif.

c, linear curve fitting plot of interaction between immobilized $Vif_{89\text{-}192}$ and A3G monomer (left), and dimer (right).

a



b


fc2-1 Vif-fl A3G Dimer 1.22-156.25nM
RU
Resp. Diff.
Time
s
A3G dimer 19.53n
A3G dimer 39.06n
A3G dimer 78.12n
A3G dimer 156.25n
A3G dimer 9.76n
A3G dimer 4.88n
A3G dimer 2.44n
A3G dimer 1.22n

c

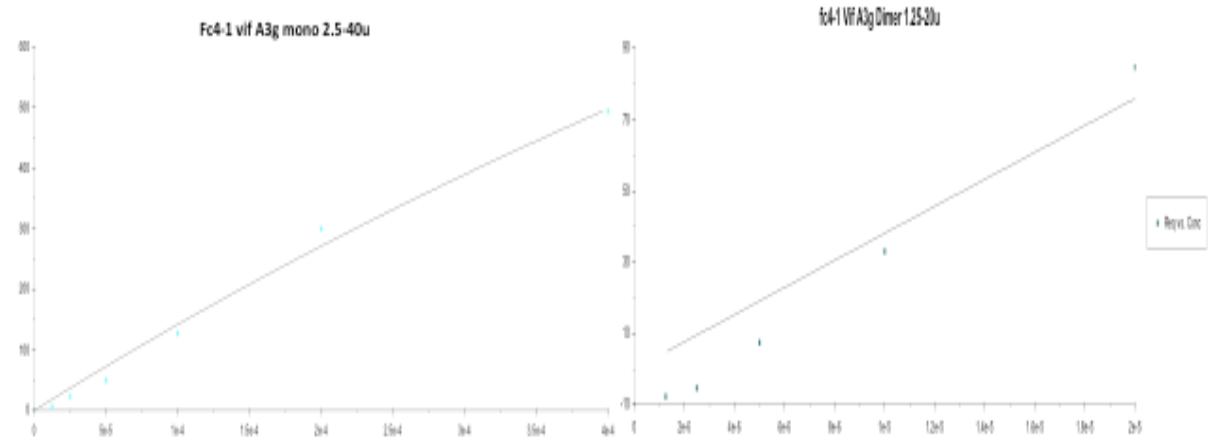

Fc4-1 vif A3g mono 2.5-40u
fc4-1 Vif A3g Dimer 1.25-20u

Table 3.2. Summary of equilibrium constant (KD) for the interaction of Vif with A3G.

KD (M)		APOBEC3G	
		Dimer	Monomer
Vif	Full-length	1.68E-10	2.05E-07
	Truncated	2.42E-03	2.03E-03

	ka (1/Ms)	kd (1/s)	KA (1/M)	KD (M)
Vif-A3G monomer	1.16E+03	2.38E-04	4.88E+06	2.05E-07
Vif-A3G dimer	4.74E+05	7.95E-05	5.96E+09	1.68E-10

CHAPTER 4

CONCLUSIONS

In recent years, the rapid advancement has been made in the studies of the biological functions of HIV-1 Vif and its interaction with human A3G. The Vif–A3G interaction is a promising potential therapeutic target of HIV-1 therapeutics, but the full potential is not likely to be realized until the detailed characterization of the Vif-A3G interaction is fully developed. Studies on the interaction between A3G and HIV-1 Vif have provided a general view of the mechanisms by which Vif mediates the antiviral activity of A3G. However, there are still many questions that remain unanswered mainly due to the lack of high-resolution three-dimensional structural and other biophysical data for HIV-1 Vif protein and full-length A3G. This lack of data is primarily due to the difficulty of producing milligram quantities of soluble recombinant Vif and A3G proteins using either prokaryotic or eukaryotic expression systems and that the purification of full-length Vif in large amounts is extremely difficult. In this study, we purified and refolded full-length Vif with a satisfactory quantity and purity for biochemical and biophysical studies. More importantly, we reported a systematic analysis of the recombinant Vif using various biophysical methodologies and provided direct evidence for the intrinsically unstructured nature of Vif under physiological conditions. Our results reveal that Vif lacks substantial secondary structure under physiological conditions as indicated by CD and NMR analyses. Vif, by itself is an intrinsically unstructured protein that has regions with poorly

defined tertiary structure when studied in solution. However, Small angle X-ray scattering analysis suggests that the Vif C-terminal region could possibly become more structured upon multimerization. Based on the studies on unbound full-length Vif, we strongly suggest that binding with its biological partners such as A3G and E3 ubiquitin ligase could also induce more order in the Vif structure. Because of the intrinsically unstructured nature of Vif, both X-ray crystallography and NMR are not very likely to be successful in providing a three-dimensional structure for Vif. Therefore, structure determination using Vif complexed with A3G, Cul5, elongin C or antibodies, would be a more promising way to obtain the high-resolution structural data.

It was important to use biochemically functional recombinant Vif in these structural studies. However, current knowledge indicates that Vif has no established enzymatic activity. Instead, it functions primarily, if not exclusively, as an adaptor molecule to mediate the interaction of other viral or host factors. In addition, we also expressed and purified the primary Vif binding partner – human A3G to test the biochemical functionality of Vif and to study Vif-A3G interaction. Our results provided direct evidence for the interaction between full-length HIV-1 Vif and A3G using biochemical and biophysical methods including AlphaScreen and surface plasmon resonance (Biacore). The A3G dimers purified as part of these studies provides the first direct evidence for dimerization of A3G under physiological conditions. By comparing the binding kinetics of A3G monomer and dimer with Vif, we found that A3G dimer has a binding affinity almost 1000 times greater than that of A3G monomer. Therefore, we proposed that although Vif can bind to both A3G monomer and dimer, dimeric A3G is the preferred conformation for Vif binding. This study also confirmed that the Vif N-terminal region accommodates the A3G binding sites by comparing the binding properties of two forms of Vif,

full-length Vif and truncated Vif89-192 variant that lacks the N-terminal A3G binding domain. Our results also revealed that both full-length and truncated recombinant Vif proteins had a strong tendency to form multimers. However, the two proteins showed distinct multimerization patterns. Unlike full-length Vif, the truncated Vif89-192 protein was found to form highly stable high-order (more than 10 mers) multimers in solution. Based on these results, we speculate that the Vif N-terminal sequence inhibits high-order Vif self-association. The mechanism that causes the differential multimerization is worth further investigation.

REFERENCES

CHAPTER 1

1. Barre-Sinoussi, F., et al., Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science, 1983. **220**(4599): p. 868-71.

2. Epidemiologic aspects of the current outbreak of Kaposi's sarcoma and opportunistic infections. N Engl J Med, 1982. **306**(4): p. 248-52.

3. Russi, E., et al., Pneumocystis carinii pneumonia and mucosal candidiasis in a previously healthy homosexual man. Infection, 1983. **11**(4): p. 196-7.

4. Gottlieb, M.S., et al., Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men: evidence of a new acquired cellular immunodeficiency. N Engl J Med, 1981. **305**(24): p. 1425-31.

5. Masur, H., et al., An outbreak of community-acquired Pneumocystis carinii pneumonia: initial manifestation of cellular immune dysfunction. N Engl J Med, 1981. **305**(24): p. 1431-8.

6. Lackner, A.A. and R.S. Veazey, Current Concepts in AIDS Pathogenesis: Insights from the SIV/Macaque Model. Annu Rev Med, 2007. **58**(1): p. 461-476.

7. Rotheram-Borus, M.J., D. Swendeman, and G. Chovnick, The Past, Present, and Future of HIV Prevention: Integrating Behavioral, Biomedical, and Structural Intervention Strategies for the Next Generation of HIV Prevention. Annual Review of Clinical Psychology, 2009. **5**(1): p. 143-167.

8. UNAIDS., Global report: UNAIDS report on the global AIDS epidemic 20102010, Geneva: UNAIDS. 359 p.

9. Gelderblom, H.R., M. Ozel, and G. Pauli, Morphogenesis and morphology of HIV. Structure-function relations. Arch Virol, 1989. **106**(1-2): p. 1-13.

10. Wu, Z., et al., Total chemical synthesis of N-myristoylated HIV-1 matrix protein p17: structural and mechanistic implications of p17 myristoylation. Proc Natl Acad Sci U S A, 2004. **101**(32): p. 11587-92.

11. Zhu, P., et al., Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature, 2006. **441**(7095): p. 847-52.

12. Zhu, P., et al., Cryoelectron tomography of HIV-1 envelope spikes: further evidence for tripod-like legs. PLoS Pathog, 2008. **4**(11): p. e1000203.

13. Freed, E.O., HIV-1 replication. Somatic cell and molecular genetics, 2001. **26**(1-6): p. 13-33.

14. Liu, R., et al., Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell, 1996. **86**(3): p. 367-77.

15. Moore, J.P., et al., The CCR5 and CXCR4 coreceptors--central to understanding the transmission and pathogenesis of human immunodeficiency virus type 1 infection. AIDS Res Hum Retroviruses, 2004. **20**(1): p. 111-26.

16. Chan, D.C. and P.S. Kim, HIV entry and its inhibition. Cell, 1998. **93**(5): p. 681-4.

17. Hiscott, J., H. Kwon, and P. Genin, Hostile takeovers: viral appropriation of the NF-kappaB pathway. J Clin Invest, 2001. **107**(2): p. 143-51.

18. Ranki, A., et al., Expression kinetics and subcellular localization of HIV-1 regulatory proteins Nef, Tat and Rev in acutely and chronically infected lymphoid cell lines. Arch Virol, 1994. **139**(3-4): p. 365-78.

19. Kim, S.Y., et al., Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression. J Virol, 1989. **63**(9): p. 3708-13.

20. Davis, A.J., P. Li, and C.J. Burrell, Kinetics of viral RNA synthesis following cell-to-cell transmission of human immunodeficiency virus type 1. J Gen Virol, 1997. **78 (Pt 8)**: p. 1897-906.

21. Ganser-Pornillos, B.K., M. Yeager, and W.I. Sundquist, The structural biology of HIV assembly. Current opinion in structural biology, 2008. **18**(2): p. 203-17.

22. Varmus, H., Retroviruses. Science, 1988. **240**(4858): p. 1427-35.

23. Greene, W.C., The molecular biology of human immunodeficiency virus type 1 infection. N Engl J Med, 1991. **324**(5): p. 308-17.

24. Adamson, C.S. and E.O. Freed, Novel approaches to inhibiting HIV-1 replication. Antiviral Res, 2010. **85**(1): p. 119-41.

25. Cullen, B.R., Mechanism of action of regulatory proteins encoded by complex retroviruses. Microbiol Rev, 1992. **56**(3): p. 375-94.

26. Sodroski, J., et al., Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science, 1985. **229**(4708): p. 74-7.

27. Somasundaran, M. and H.L. Robinson, Unexpectedly high levels of HIV-1 RNA and protein synthesis in a cytocidal infection. Science, 1988. **242**(4885): p. 1554-7.

28. Arya, S.K., et al., Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). Science, 1985. **229**(4708): p. 69-73.

29. Wei, P., et al., A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell, 1998. **92**(4): p. 451-62.

30. Zhou, M., et al., The Tat/TAR-dependent phosphorylation of RNA polymerase II C-terminal domain stimulates cotranscriptional capping of HIV-1 mRNA. Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12666-71.

31. Mancebo, H.S., et al., P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes & development, 1997. **11**(20): p. 2633-44.

32. Frank, P., HIV Nef: The Mother of All Evil? Immunity, 1998. **9**(4): p. 433-437.

33. Greenway, A.L., G. Holloway, and D.A. McPhee, HIV-1 Nef: a critical factor in viral-induced pathogenesis. Advances in pharmacology, 2000. **48**: p. 299-343.

34. Tobiume, M., et al., Inefficient enhancement of viral infectivity and CD4 downregulation by human immunodeficiency virus type 1 Nef from Japanese long-term nonprogressors. J Virol, 2002. **76**(12): p. 5959-65.

35. Deacon, N.J., et al., Genomic structure of an attenuated quasi species of HIV-1 from a blood transfusion donor and recipients. Science, 1995. **270**(5238): p. 988-91.

36. Stoddart, C.A., et al., Human immunodeficiency virus type 1 Nef-mediated downregulation of CD4 correlates with Nef enhancement of viral pathogenesis. J Virol, 2003. **77**(3): p. 2124-33.

37. Wonderlich, E.R., et al., ARF-1 activity is required to recruit AP-1 to the MHC-I cytoplasmic tail and disrupt MHC-I trafficking in HIV-1 infected primary T cells. J Virol, 2011.

38. Tokarev, A. and J. Guatelli, Misdirection of membrane trafficking by HIV-1 Vpu and Nef: Keys to viral virulence and persistence. Cellular logistics, 2011. **1**(3): p. 90-102.

39. Jabbar, M.A. and D.P. Nayak, Intracellular interaction of human immunodeficiency virus type 1 (ARV-2) envelope glycoprotein gp160 with CD4 blocks the movement and maturation of CD4 to the plasma membrane. J Virol, 1990. **64**(12): p. 6297-304.

40. Klimkait, T., et al., The human immunodeficiency virus type 1-specific protein vpu is required for efficient virus maturation and release. J Virol, 1990. **64**(2): p. 621-9.

41. Kagi, D., et al., Molecular mechanisms of lymphocyte-mediated cytotoxicity and their role in immunological protection and pathogenesis in vivo. Annu Rev Immunol, 1996. **14**: p. 207-32.

42. Collins, K.L., et al., HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature, 1998. **391**(6665): p. 397-401.

43. Lindwasser, O.W., R. Chaudhuri, and J.S. Bonifacino, Mechanisms of CD4 downregulation by the Nef and Vpu proteins of primate immunodeficiency viruses. Curr Mol Med, 2007. **7**(2): p. 171-84.

44. Burtey, A., et al., Dynamic interaction of HIV-1 Nef with the clathrin-mediated endocytic pathway at the plasma membrane. Traffic, 2007. **8**(1): p. 61-76.

45. Bentham, M., S. Mazaleyrat, and M. Harris, The di-leucine motif in the cytoplasmic tail of CD4 is not required for binding to human immunodeficiency virus type 1 Nef, but is critical for CD4 down-modulation. J Gen Virol, 2003. **84**(Pt 10): p. 2705-13.

46. Schaefer, M.R., et al., HIV-1 Nef targets MHC-I and CD4 for degradation via a final common beta-COP-dependent pathway in T cells. PLoS Pathog, 2008. **4**(8): p. e1000131.

47. Strebel, K., Virus-host interactions: role of HIV proteins Vif, Tat, and Rev. AIDS, 2003. **17 Suppl 4**: p. S25-34.

48. Pollard, V.W. and M.H. Malim, The HIV-1 Rev protein. Annual Review of Microbiology, 1998. **52**: p. 491-532.

49. Robertson-Anderson, R.M., et al., Single-molecule studies reveal that DEAD box protein DDX1 promotes oligomerization of HIV-1 Rev on the Rev response element. J Mol Biol, 2011. **410**(5): p. 959-71.

50. Checkley, M.A., B.G. Luttge, and E.O. Freed, HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J Mol Biol, 2011. **410**(4): p. 582-608.

51. Freed, E.O., HIV-1 gag proteins: diverse functions in the virus life cycle. Virology, 1998. **251**(1): p. 1-15.

52. Jouvenet, N., et al., Plasma membrane is the site of productive HIV-1 particle assembly. PLoS biology, 2006. **4**(12): p. e435.

53. Resh, M.D., A myristoyl switch regulates membrane binding of HIV-1 Gag. Proc Natl Acad Sci U S A, 2004. **101**(2): p. 417-8.

54. Kutluay, S.B. and P.D. Bieniasz, Analysis of the initiating events in HIV-1 particle assembly and genome packaging. PLoS Pathog, 2010. **6**(11): p. e1001200.

55. Adamson, C.S. and I.M. Jones, The molecular basis of HIV capsid assembly--five years of progress. Rev Med Virol, 2004. **14**(2): p. 107-21.

56. Domenech, R., et al., The isolated major homology region of the HIV capsid protein is mainly unfolded in solution and binds to the intact protein. Biochim Biophys Acta, 2011. **1814**(10): p. 1269-78.

57. Lever, A., et al., Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J Virol, 1989. **63**(9): p. 4085-7.

58. De Guzman, R.N., et al., Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element. Science, 1998. **279**(5349): p. 384-8.

59. Ohishi, M., et al., The relationship between HIV-1 genome RNA dimerization, virion maturation and infectivity. Nucleic Acids Res, 2011. **39**(8): p. 3404-17.

60. Demirov, D.G., J.M. Orenstein, and E.O. Freed, The late domain of human immunodeficiency virus type 1 p6 promotes virus release in a cell type-dependent manner. J Virol, 2002. **76**(1): p. 105-17.

61. Gottlinger, H.G., et al., Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc Natl Acad Sci U S A, 1991. **88**(8): p. 3195-9.

62. Gottwein, E., et al., Cumulative mutations of ubiquitin acceptor sites in human immunodeficiency virus type 1 gag cause a late budding defect. J Virol, 2006. **80**(13): p. 6267-75.

63. Demirov, D.G., et al., Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. Proc Natl Acad Sci U S A, 2002. **99**(2): p. 955-60.

64. Yeager, M., et al., Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: implications for retroviral assembly mechanisms. Proc Natl Acad Sci U S A, 1998. **95**(13): p. 7299-304.

65. Datta, S.A., et al., On the role of the SP1 domain in HIV-1 particle assembly: a molecular switch? J Virol, 2011. **85**(9): p. 4111-21.

66. Bukrinskaya, A., HIV-1 matrix protein: a mysterious regulator of the viral life cycle. Virus Res, 2007. **124**(1-2): p. 1-11.

67. Wensing, A.M., N.M. van Maarseveen, and M. Nijhuis, Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. Antiviral Res, 2010. **85**(1): p. 59-74.

68. Malim, M.H. and M. Emerman, HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe, 2008. **3**(6): p. 388-98.

69. Ruiz, A., J.C. Guatelli, and E.B. Stephens, The Vpu protein: new concepts in virus release and CD4 down-modulation. Curr HIV Res, 2010. **8**(3): p. 240-52.

70. Bour, S., U. Schubert, and K. Strebel, The human immunodeficiency virus type 1 Vpu protein specifically binds to the cytoplasmic domain of CD4: implications for the mechanism of degradation. J Virol, 1995. **69**(3): p. 1510-20.

71. Estrabaud, E., et al., Regulated degradation of the HIV-1 Vpu protein through a betaTrCP-independent pathway limits the release of viral particles. PLoS Pathog, 2007. **3**(7): p. e104.

72. Terwilliger, E.F., et al., Functional role of human immunodeficiency virus type 1 vpu. Proc Natl Acad Sci U S A, 1989. **86**(13): p. 5163-7.

73. Strebel, K., et al., Molecular and biochemical analyses of human immunodeficiency virus type 1 vpu protein. J Virol, 1989. **63**(9): p. 3784-91.

74. Varthakavi, V., et al., Viral protein U counteracts a human host cell restriction that inhibits HIV-1 particle production. Proc Natl Acad Sci U S A, 2003. **100**(25): p. 15154-9.

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

76. Nomaguchi, M., M. Fujita, and A. Adachi, Role of HIV-1 Vpu protein for virus spread and pathogenesis. Microbes Infect, 2008. **10**(9): p. 960-7.

77. Planelles, V. and S. Benichou, Vpr and its interactions with cellular proteins. Current topics in microbiology and immunology, 2009. **339**: p. 177-200.

78. Morellet, N., et al., NMR structure of the HIV-1 regulatory protein VPR. J Mol Biol, 2003. **327**(1): p. 215-27.

79. Zhao, R.Y., G. Li, and M.I. Bukrinsky, Vpr-host interactions during HIV-1 viral life cycle. J Neuroimmune Pharmacol, 2011. **6**(2): p. 216-29.

80. He, J., et al., Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. J Virol, 1995. **69**(11): p. 6705-11.

81. Davy, C. and J. Doorbar, G2/M cell cycle arrest in the life cycle of viruses. Virology, 2007. **368**(2): p. 219-26.

82. Goh, W.C., et al., HIV-1 Vpr increases viral expression by manipulation of the cell cycle: a mechanism for selection of Vpr in vivo. Nat Med, 1998. **4**(1): p. 65-71.

83. Ehrlich, E.S. and X.F. Yu, Lentiviral Vif: viral hijacker of the ubiquitin-proteasome system. Int J Hematol, 2006. **83**(3): p. 208-12.

84. Chiu, Y.L. and W.C. Greene, APOBEC3G: an intracellular centurion. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2009. **364**(1517): p. 689-703.

85. Chiu, Y.-L. and W.C. Greene, The APOBEC3 Cytidine Deaminases: An Innate Defensive Network Opposing Exogenous Retroviruses and Endogenous Retroelements. Annu Rev Immunol, 2008. **26**(1): p. 317-353.

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

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

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

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

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

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

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

93. Malim, M.H., APOBEC proteins and intrinsic resistance to HIV-1 infection. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2009. **364**(1517): p. 675-87.

94. Wissing, S., N.L. Galloway, and W.C. Greene, HIV-1 Vif versus the APOBEC3 cytidine deaminases: an intracellular duel between pathogen and host restriction factors. Mol Aspects Med, 2010. **31**(5): p. 383-97.

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

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

97. Knoepfel, S.A., et al., Comparison of G-to-A mutation frequencies induced by APOBEC3 proteins in H9 cells and peripheral blood mononuclear cells in the context of impaired processivities of drug-resistant human immunodeficiency virus type 1 reverse transcriptase variants. J Virol, 2008. **82**(13): p. 6536-45.

98. Thielen, B.K., et al., T cells contain an RNase-insensitive inhibitor of APOBEC3G deaminase activity. PLoS Pathog, 2007. **3**(9): p. 1320-34.

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

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

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

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

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

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

105. Chen, K.M., et al., Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. Nature, 2008. **452**(7183): p. 116-9.

106. Bransteitter, R., C. Prochnow, and X.S. Chen, The current structural and functional understanding of APOBEC deaminases. Cellular and molecular life sciences : CMLS, 2009. **66**(19): p. 3137-47.

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

108. Barraud, P., et al., Advances in the structural understanding of Vif proteins. Curr HIV Res, 2008. **6**(2): p. 91-9.

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

110. Yu, Y., et al., Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. Genes Dev, 2004. **18**(23): p. 2867-72.

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

112. Walker, R.C., Jr., et al., Identification of dominant negative human immunodeficiency virus type 1 Vif mutants that interfere with the functional inactivation of APOBEC3G by virus-encoded Vif. J Virol, 2010. **84**(10): p. 5201-11.

113. Yang, S., Y. Sun, and H. Zhang, The multimerization of human immunodeficiency virus type I Vif protein: a requirement for Vif function in the viral life cycle. J Biol Chem, 2001. **276**(7): p. 4889-93.

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

115. Yang, B., et al., Potent suppression of viral infectivity by the peptides that inhibit multimerization of human immunodeficiency virus type 1 (HIV-1) Vif proteins. J Biol Chem, 2003. **278**(8): p. 6596-602.

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

117. Adessi, C. and C. Soto, Converting a peptide into a drug: strategies to improve stability and bioavailability. Curr Med Chem, 2002. **9**(9): p. 963-78.

118. Smith, J.L., et al., Multiple ways of targeting APOBEC3-virion infectivity factor interactions for anti-HIV-1 drug development. Trends in pharmacological sciences, 2009. **30**(12): p. 638-46.

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

CHAPTER 2

1. Frankel, A.D. and J.A. Young, HIV-1: fifteen proteins and an RNA. Annual Review of Biochemistry, 1998. 67: p. 1-25.

2. Freed, E.O., HIV-1 replication. Somatic cell and molecular genetics, 2001. 26(1-6): p. 13-33.

3. Gramberg, T., N. Sunseri, and N.R. Landau, Accessories to the crime: recent advances in HIV accessory protein biology. Curr HIV/AIDS Rep, 2009. 6(1): p. 36-42.

4. Pollard, V.W. and M.H. Malim, The HIV-1 Rev protein. Annual Review of Microbiology, 1998. 52: p. 491-532.

5. Lindwasser, O.W., R. Chaudhuri, and J.S. Bonifacino, Mechanisms of CD4 downregulation by the Nef and Vpu proteins of primate immunodeficiency viruses. Curr Mol Med, 2007. 7(2): p. 171-84.

6. Bukrinsky, M. and A. Adzhubei, Viral protein R of HIV-1. Rev Med Virol, 1999. 9(1): p. 39-49.

7. Le Rouzic, E. and S. Benichou, The Vpr protein from HIV-1: distinct roles along the viral life cycle. Retrovirology, 2005. 2: p. 11.

8. Planelles, V. and S. Benichou, Vpr and its interactions with cellular proteins. Current topics in microbiology and immunology, 2009. 339: p. 177-200.

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

10. Ruiz, A., J.C. Guatelli, and E.B. Stephens, The Vpu protein: new concepts in virus release and CD4 down-modulation. Curr HIV Res, 2010. 8(3): p. 240-52.

11. Malim, M.H. and M. Emerman, HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe, 2008. 3(6): p. 388-98.

12. Ehrlich, E.S. and X.F. Yu, Lentiviral Vif: viral hijacker of the ubiquitin-proteasome system. Int J Hematol, 2006. 83(3): p. 208-12.

13. Chiu, Y.L. and W.C. Greene, APOBEC3G: an intracellular centurion. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2009. 364(1517): p. 689-703.

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

15. Chiu, Y.-L. and W.C. Greene, The APOBEC3 Cytidine Deaminases: An Innate Defensive Network Opposing Exogenous Retroviruses and Endogenous Retroelements. Annu Rev Immunol, 2008. 26(1): p. 317-353.

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

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

18. Reingewertz, T.H., et al., The C-terminal domain of the HIV-1 Vif protein is natively unfolded in its unbound state. Protein Eng Des Sel, 2009. 22(5): p. 281-7.

19. Reingewertz, T.H., D.E. Shalev, and A. Friedler, Structural disorder in the HIV-1 Vif protein and interaction-dependent gain of structure. Protein Pept Lett, 2010. 17(8): p. 988-98.

20. Marcsisin, S.R., et al., On the Solution Conformation and Dynamics of the HIV-1 Viral Infectivity Factor. J Mol Biol, 2011. 410(5): p. 1008-22.

21. Auclair, J.R., et al., Mass spectrometry analysis of HIV-1 Vif reveals an increase in ordered structure upon oligomerization in regions necessary for viral infectivity. Proteins, 2007. 69(2): p. 270-84.

22. Sreerama, N. and R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem, 2000. 287(2): p. 252-60.

23. Whitmore, L. and B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res, 2004. 32(Web Server issue): p. W668-73.

24. Svergun, D., Determination of the regularization parameter in indirect-transform methods using perceptual criteria. Journal of Applied Crystallography, 1992. 25(4): p. 495-503.

25. Svergun, D.I., M.V. Petoukhov, and M.H. Koch, Determination of domain structure of proteins from X-ray solution scattering. Biophysical journal, 2001. 80(6): p. 2946-53.

26. Volkov, V.V. and D.I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering. Journal of Applied Crystallography, 2003. 36(3 Part 1): p. 860-864.

27. Yang, S., Y. Sun, and H. Zhang, The multimerization of human immunodeficiency virus type I Vif protein: a requirement for Vif function in the viral life cycle. J Biol Chem, 2001. 276(7): p. 4889-93.

28. Bernacchi, S., et al., Importance of the proline-rich multimerization domain on the oligomerization and nucleic acid binding properties of HIV-1 Vif. Nucleic Acids Res, 2011. 39(6): p. 2404-15.

29. Yang, B., et al., Potent suppression of viral infectivity by the peptides that inhibit multimerization of human immunodeficiency virus type 1 (HIV-1) Vif proteins. J Biol Chem, 2003. 278(8): p. 6596-602.

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

31. Donahue, J.P., et al., The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. Virology, 2008. 377(1): p. 49-53.

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

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

34. Wagner, G. and K. Wuthrich, Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra. Basic pancreatic trypsin inhibitor. J Mol Biol, 1982. 155(3): p. 347-66.

35. Ciaccio, N.A., et al., High-yield expression in E. coli and refolding of the bZIP domain of activating transcription factor 5. Protein Expr Purif, 2008. 62(2): p. 235-43.

36. Pantoliano, M.W., et al., High-density miniaturized thermal shift assays as a general strategy for drug discovery. Journal of biomolecular screening, 2001. 6(6): p. 429-40.

37. Putnam, C.D., et al., X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Quarterly reviews of biophysics, 2007. 40(3): p. 191-285.

38. Blobel, J., et al., Low-resolution structures of transient protein-protein complexes using small-angle X-ray scattering. Journal of the American Chemical Society, 2009. 131(12): p. 4378-86.

39. Rambo, R.P. and J.A. Tainer, Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. Current opinion in structural biology, 2010. 20(1): p. 128-37.

40. Kohn, J.E., et al., Random-coil behavior and the dimensions of chemically unfolded proteins. Proc Natl Acad Sci U S A, 2004. 101(34): p. 12491-6.

41. Walker, R.C., Jr., et al., Identification of dominant negative human immunodeficiency virus type 1 Vif mutants that interfere with the functional inactivation of APOBEC3G by virus-encoded Vif. J Virol, 2010. 84(10): p. 5201-11.

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

43. Li, X., et al., Predicting Protein Disorder for N-, C-, and Internal Regions. Genome informatics. Workshop on Genome Informatics, 1999. 10: p. 30-40.

44. Romero, P., et al., Sequence complexity of disordered protein. Proteins, 2001. 42(1): p. 38-48.

45. Romero, Obradovic, and K. Dunker, Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. Genome informatics. Workshop on Genome Informatics, 1997. 8: p. 110-124.

46. Barraud, P., et al., Advances in the structural understanding of Vif proteins. Curr HIV Res, 2008. 6(2): p. 91-9.

CHAPTER 3

1. Ehrlich, E.S. and X.F. Yu, Lentiviral Vif: viral hijacker of the ubiquitin-proteasome system. Int J Hematol, 2006. 83(3): p. 208-12.

2. Chiu, Y.L. and W.C. Greene, APOBEC3G: an intracellular centurion. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2009. 364(1517): p. 689-703.

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

4. Chiu, Y.-L. and W.C. Greene, The APOBEC3 Cytidine Deaminases: An Innate Defensive Network Opposing Exogenous Retroviruses and Endogenous Retroelements. Annu Rev Immunol, 2008. 26(1): p. 317-353.

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

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

7. Wissing, S., N.L. Galloway, and W.C. Greene, HIV-1 Vif versus the APOBEC3 cytidine deaminases: an intracellular duel between pathogen and host restriction factors. Mol Aspects Med, 2010. 31(5): p. 383-97.

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

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

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

11. Reingewertz, T.H., et al., The C-terminal domain of the HIV-1 Vif protein is natively unfolded in its unbound state. Protein Eng Des Sel, 2009. 22(5): p. 281-7.

12. Reingewertz, T.H., D.E. Shalev, and A. Friedler, Structural disorder in the HIV-1 Vif protein and interaction-dependent gain of structure. Protein Pept Lett, 2010. 17(8): p. 988-98.

13. Marcsisin, S.R., et al., On the Solution Conformation and Dynamics of the HIV-1 Viral Infectivity Factor. J Mol Biol, 2011. 410(5): p. 1008-22.

14. Auclair, J.R., et al., Mass spectrometry analysis of HIV-1 Vif reveals an increase in ordered structure upon oligomerization in regions necessary for viral infectivity. Proteins, 2007. 69(2): p. 270-84.

15. Yang, S., Y. Sun, and H. Zhang, The multimerization of human immunodeficiency virus type I Vif protein: a requirement for Vif function in the viral life cycle. J Biol Chem, 2001. 276(7): p. 4889-93.

16. Yang, B., et al., Potent suppression of viral infectivity by the peptides that inhibit multimerization of human immunodeficiency virus type 1 (HIV-1) Vif proteins. J Biol Chem, 2003. 278(8): p. 6596-602.

17. Donahue, J.P., et al., The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. Virology, 2008. 377(1): p. 49-53.

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

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

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

21. Wolfe, L.S., et al., Dissection of the HIV Vif interaction with human E3 ubiquitin ligase. J Virol, 2010. 84(14): p. 7135-9.

22. Bernacchi, S., et al., Importance of the proline-rich multimerization domain on the oligomerization and nucleic acid binding properties of HIV-1 Vif. Nucleic Acids Res, 2011. 39(6): p. 2404-15.

23. Pollard, V.W. and M.H. Malim, The HIV-1 Rev protein. Annual Review of Microbiology, 1998. 52: p. 491-532.

24. Kwak, Y.T., et al., Self-association of the Lentivirus protein, Nef. Retrovirology, 2010. 7: p. 77.

25. Maldarelli, F., et al., Human immunodeficiency virus type 1 Vpu protein is an oligomeric type I integral membrane protein. J Virol, 1993. 67(8): p. 5056-61.

26. Vigan, R. and S.J. Neil, Determinants of tetherin antagonism in the transmembrane domain of the human immunodeficiency virus type 1 Vpu protein. J Virol, 2010. 84(24): p. 12958-70.

27. Venkatachari, N.J., et al., Human immunodeficiency virus type 1 Vpr: oligomerization is an essential feature for its incorporation into virus particles. Virol J, 2010. 7: p. 119.

28. Frankel, A.D., D.S. Bredt, and C.O. Pabo, Tat protein from human immunodeficiency virus forms a metal-linked dimer. Science, 1988. 240(4848): p. 70-3.

29. Bogerd, H.P., et al., Genetic evidence that the Tat proteins of human immunodeficiency virus types 1 and 2 can multimerize in the eukaryotic cell nucleus. J Virol, 1993. 67(8): p. 5030-4.

30. Hartlieb, B., et al., Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. J Biol Chem, 2003. 278(43): p. 41830-6.

31. Nemeroff, M.E., X.Y. Qian, and R.M. Krug, The influenza virus NS1 protein forms multimers in vitro and in vivo. Virology, 1995. 212(2): p. 422-8.

32. Brindley, M.A. and R.K. Plemper, Blue native PAGE and biomolecular complementation reveal a tetrameric or higher-order oligomer organization of the physiological measles virus attachment protein H. J Virol, 2010. 84(23): p. 12174-84.

33. Zamoto-Niikura, A., et al., Rift valley fever virus L protein forms a biologically active oligomer. J Virol, 2009. 83(24): p. 12779-89.

34. Rozzelle, J.E., Jr., et al., Self-association of a synthetic peptide from the N terminus of the hepatitis delta virus protein into an immunoreactive alpha-helical multimer. Proc Natl Acad Sci U S A, 1995. 92(2): p. 382-6.

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

36. Chiu, Y.L. and W.C. Greene, APOBEC3G: an intracellular centurion. Philos Trans R Soc Lond B Biol Sci, 2009. 364(1517): p. 689-703.

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

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

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

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

41. Salter, J.D., et al., A hydrodynamic analysis of APOBEC3G reveals a monomer-dimer-tetramer self-association that has implications for anti-HIV function. Biochemistry, 2009. 48(45): p. 10685-7.

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

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

44. McDougall, W.M., C. Okany, and H.C. Smith, Deaminase activity on ssDNA occurred in vitro when APOBEC3G forms homotetramers and higher-order complexes. J Biol Chem, 2011.

45. Bulenger, S., S. Marullo, and M. Bouvier, Emerging role of homo- and heterodimerization in G-protein-coupled receptor biosynthesis and maturation. Trends in pharmacological sciences, 2005. 26(3): p. 131-7.

46. Li, E. and K. Hristova, Receptor tyrosine kinase transmembrane domains: Function, dimer structure and dimerization energetics. Cell adhesion & migration, 2010. 4(2): p. 249-54.

47. Dutzler, R., et al., X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. Nature, 2002. 415(6869): p. 287-94.

48. Ispolatov, I., et al., Binding properties and evolution of homodimers in protein-protein interaction networks. Nucleic Acids Res, 2005. 33(11): p. 3629-35.

49. Marianayagam, N.J., M. Sunde, and J.M. Matthews, The power of two: protein dimerization in biology. Trends Biochem Sci, 2004. 29(11): p. 618-25.

50. Quinn, A.M., et al., A homogeneous method for investigation of methylation-dependent protein-protein interactions in epigenetics. Nucleic Acids Res, 2010. 38(2): p. e11.

51. Myszka, D.G., Improving biosensor analysis. Journal of molecular recognition : JMR, 1999. 12(5): p. 279-84.

52. Abdiche, Y.N., D.S. Malashock, and J. Pons, Probing the binding mechanism and affinity of tanezumab, a recombinant humanized anti-NGF monoclonal antibody, using a repertoire of biosensors. Protein science : a publication of the Protein Society, 2008. 17(8): p. 1326-35.

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

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

55. Walker, R.C., Jr., et al., Identification of dominant negative human immunodeficiency virus type 1 Vif mutants that interfere with the functional inactivation of APOBEC3G by virus-encoded Vif. J Virol, 2010. 84(10): p. 5201-11.

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

57. Li, X., et al., Predicting Protein Disorder for N-, C-, and Internal Regions. Genome informatics. Workshop on Genome Informatics, 1999. 10: p. 30-40.

58. Romero, P., et al., Sequence complexity of disordered protein. Proteins, 2001. 42(1): p. 38-48.

59. Romero, Obradovic, and K. Dunker, Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. Genome informatics. Workshop on Genome Informatics, 1997. 8: p. 110-124.

60. Barraud, P., et al., Advances in the structural understanding of Vif proteins. Curr HIV Res, 2008. 6(2): p. 91-9.

APPENDIX

PRODUCTION OF EBOLA VIRUS SECRETORY GLYCOPROTEIN (SGP) USING BACTERIAL, BACULOVIRUS AND MAMMALIAN EXPRESSION SYSTEMS[3]

[3]Zhou, D., Xu, H., Wang, B.C., Hogan., J. and Rose, J.
To be submitted to *The Journal of Biological Chemistry*

ABSTRACT

Ebola virus, being highly pathogenic for humans and a potential subject of biological weaponization, is now one of the most feared and dangerous pathogens worldwide [1]. Currently, there is no therapeutic agents or vaccines available for human use against Ebola virus, not even treatments useful shortly after virus exposure. Therefore, the only available clinical option is supportive care. Extensive efforts over the last 30 years have been made to develop vaccines against Ebola virus with little success. Classical approaches have been generally unsuccessful and pose enormous safety concerns. Modern approaches, especially vector-based approaches, have achieved success in nonhuman primate models [2, 3]. Despite these successes, more studies remains to be done. Of the progress that has been made, the most promising vaccine candidates share a common theme: induction of a host immune response against the Ebola virus glycoprotein (GP). However, protein production and post-translational modifications in prokaryotic and eukaryotic systems differ greatly. Thus, in terms of effective vaccine development, it's extremely important to understand whether the source of the recombinant protein used to produce the vaccine would affect the protein structure and subsequent antigen-specific immune response. To answer this question, in this study we expressed and purified Ebola secretory glycoprotein (sGP) from three independent expression systems – *E. coli*, insect and mammalian cells for the generation of antibodies to sGP and examination of the immune responses of the host animals. Focused on using structural biology and biophysical methodologies to characterize the sGP proteins obtained from different expression systems, the correlation between the protein production platforms and host immune responses will be characterized ultimately. The results of this study will have direct applicability to vaccine

production and aid in our long-term goal of using structural biology to aid in the design of more effective vaccines.

INTRODUCTION

The Ebola virus (EBOV) is a highly pathogenic virus in human and non-human primates. EBOV is one of the most virulent pathogens with a mortality that can be as high as 90% and has an extremely rapid disease course that can cause death within days of infection [4]. Ebola virus, first discovered in 1976 in Africa, has received considerable attention in recent years due to its increasing re-emergence in Africa and its potential for use as a biological weapon. Ebola virus is currently subdivided into four distinct species: the Zaire (ZEBOV), Sudan (SEBOV), Cote d'Ivoire (CIEBOV), and Reston Ebola virus (REBOV). Among them, ZEBOV was thought to be the most virulent of all Ebola viruses, with a mortality around 80% [5]. Ebola viruses as a group are classified as "Category A" biological weapons by the US Centers for Disease Control and Prevention (CDC, www.cdc.gov) because of their great potential for adversely impacting public health due to the mass casualties resulting from infection [6]. However, there is no vaccine or treatment available for human use.

Ebola viruses are enveloped single-stranded negative-sense linear RNA viruses with an unusual filamentous morphology sometimes called the Shepard's crook. The RNA genome, approximately 19-kb long, is non-infectious and does not contain a poly(A) tail. After entry into the cytoplasm of host cells, it is transcribed to generate polyadenylated subgenomic mRNA encoding seven viral genes: (1) NP (major nucleoprotein), (2) VP35 (P-like protein), (3) VP40 (matrix protein), (4) GP (glycoprotein), (5) VP30 (minor nucleoprotein), (6) VP24, and (7) L (RNA-dependent RNA polymerase). The viral genes, gene order, and presumptive protein functions are shown in Figure A.1. Of these seven proteins of Ebola virus, four proteins (NP, VP30, VP35, and L) are associated with the virus genomic RNA in the ribonucleoprotein

complex, while the remaining three structural proteins (GP, VP40 and VP24) are associated with the viral membrane [5].

Immune response to Ebola virus

It is becoming increasingly clear that EBOV overwhelms host immune defenses by initially dysregulating and defeating innate immune system followed by a similar disruption of the adaptive immune system of both primates and humans. In primates and humans, the most distinguishing feature of EBOV infection is unchecked viral replication that outpaces the protein synthesis and host immune response in infected organisms [7]. In addition, EBOV infection exhibits a relatively wide range of symptoms including bleeding, fever, malaise, diarrhea, vomiting, and severe liver damage. Extensive studies have shown that after entry into host, EBOV targets cells of the monocyte/macrophage lineage, resulting in the release of several pro-inflammatory cytokines such as interferon γ (IFNγ) and tumor-necrosis factor (TNF). The release of cytokines then leads to damage of the vascular endothelium [4, 7-10]. In addition to the damage to vascular permeability caused by cytokine effects the excessive bleeding associated with EBOV infection can also be attributed to a decrease in platelet numbers, severe liver damage and the activation of tissue factors in monocytes and macrophages [11, 12].

Vaccine development

There has been a long-standing debate about the necessity of vaccine for EBOV. The rare occurrence of EBOV and the remote location of the outbreaks did not warrant the cost of vaccine development. However, this view has changed now that EBOV has become a potential tool for bioterrorism [1, 13]. Historically, EBOV vaccine development began shortly after the discovery

of the virus in 1976. The earliest attempts were based on the classical approach; *i.e.* formalin-fixed or heat-inactivated virus. Classical attenuation by passage through cell culture or non-human primates was also attempted for vaccine development [2, 3]. Although, important advances have been made in developing preventive vaccines that are protective in animal models, no effective therapies and/or vaccines are available for human use.

More recently EBOV vaccine development has been focusing on the use of subunit vaccines that are based on a single or combinations of Ebola-encoded structural proteins to induce protective immunity. The EBOV glycoprotein (GP), nucleoprotein (NP), and the virus structural proteins VP24, VP30, VP35 and VP40 have been tested as vaccine candidates using naked DNA, adenovirus and virus-like replicon particles as delivery mechanisms [4]. In 2004, a DNA vaccine expressing the EBOV GP was tested in phase I clinical trials [14]. EBOV GP is gaining increased attention for the vaccine development since the protein is exposed on the surface of the virion and appears to be the only target for the neutralizing antibody [15-19]. Previous studies showed that neutralizing antibodies are able to inhibit viral infection in culture and prevent viral infection in mice when present prior to infection at sufficient concentration. These EBOV neutralizing antibodies can be provided through passive administration or induced by vaccination [15-19]. The work presented here focuses production of EBOV soluble glycoprotein sGP using three common protein production platforms (bacterial expression, baculorvirus expression and mammalian cell expression) related subunit based vaccine production.

The EBOV *gp* gene encodes for the precursors of two glycoproteins, the soluble nonstructural glycoprotein (pre-sGP) and the transmembrane structural glycoprotein (pre-GP). Pre-sGP is the primary gene product and is processed by furin cleavage into the 364-residue sGP

protein that is secreted in high levels from infected cells [20, 21]. By contrast, pre-GP and post-translationally cleaved by furin into a large 500 residue amino-terminal fragment (GP1) and a smaller 179-reside carboxyl-terminal fragment (GP2). GP1 and GP2 associate by disulfide bonds, and the mature GP1-GP2 complex is anchored in membrane by a transmembrane domain located near the carboxyl terminus of GP2 [22]. Although, sGP and GP1 share the same N-terminal sequence (295 amino acids), the two proteins have distinct biochemical and biological properties. For example, GP appears to form a trimeric complex and binds preferentially to endothelial cells [23], whereas sGP forms a dimeric (head-to-head) complex (sGP_2) that binds to neutrophil cells [24]. These GP proteins also differ in their functions; GP1 and GP2 have been shown to mediate virus entry, attachment and fusion [22], whereas sGP is not involved in virus entry but is thought to have an anti-inflammatory role [9]. In addition, while both GP1 and GP2 are heavily N- and O-linked glycosylated, sGP exhibits primarily N-linked glycosylation. sGP also exhibits C-mannosylation [25, 26] and sialylated glycans wshich are not generally observed in GP1 and GP2 [27].

As noted above, EBOV GP proteins (GP1 and GP2) are highly glycosylated with approximately half of their apparent molecular weight comprised of N-linked and O-linked carbohydrates. Studies of the GP from different Ebola virus strains indicate at least 17 possible sites of N-linked glycosylation, and multiple additional sites of O-linked glycosylation within the mucin region of EBOV GP [28]. Yet, the physiological role of N-linked or O-linked glycosylation of the EBOV GP has received very little attention. It has been demonstrated that glycosylation can greatly influence the structure, function, antigenicity, and immunogenicity of various viral glycoproteins [29]. For example, extensive glycosylation has been found on the envelope protein - gp120 of human immunodeficiency virus (HIV) and is involved in the

structural integrity, antigenicity, and immunogenicity of the protein. The gp120 carbohydrates are clustered together on an outer shell of the protein and have been suggested to form a "glycan shield" that prevents neutralizing antibodies from binding due to steric hindrance [30]. In addition, the glycosylation of viral structural proteins might also prevent the induction of antibodies in the early stage of viral infection [31]. A recent study from Hogan *et al.* showed that: (1) removal of potential glycosylation sites of EBOV GP resulted in altered antibody and T cell responses; and (2) extensive glycosylation of EBOV GP prevented robust humoral and cell-mediated immune responses that might otherwise provide protection [29].

With this in mind, we hypothesized that Ebola GP proteins produced from different protein production systems- *E.coli*, insect/baculovirus and mammalian cell, might induce different humoral and cell mediated immune responses, because proteins produced in *E. coli* have no post-translational glycosylation; while insect cells exhibit altered glycosylation pattern from that of mammalian cells [32]. Here we report the successful expression and purification of EBOV sGP using three independent expression systems – *E. coli*, insect and mammalian cells needed for antibody generation and subsequent examination of their immune responses in host animals.

EXPERIMENTAL METHODS

Cloning and vector construction

The EBOV (Zaire) sGP construct was generated by PCR using pcDNA3.1 vector containing full-length EBOV sGP cDNA (provided from Dr. Jeff Hogan at University of Georgia) as the template. The 5' and 3' primer sequence were

GAAAACCTGTACTTCCAAGGCGGGTCAGGTGGTGTTACAGGAATATTGC and GGGGACCACTTTGTACAAGAAAGCTGGGTTTAGCGACACTGCAGCTT, respectively. The PCR product was purified by agarose gel electrophoresis and isolated using a QIAquick Gel Extraction kit (QIAGEN), followed by cloning into the pDONR221 vector using BP Clonase (Invitrogen) to generate the Gateway® entry clone. The resulting plasmid, pDONR221-sGP was used as the basis for expression plasmid construction. Site-specific recombination was then carried out to subclone the sGP gene into the Gateway® destination vectors – pDEST527 (for *E. coli* expression), pDEST10 (for baculorvirus expression) and pDEST26 (for mammalian expression), see Table A.1. All constructs contained a N-terminal 6×His purification tag followed by a TEV protease cleavage site to aid in purification. To generate the expression vectors – pDEST527-sGP, pDEST10-sGP and pDEST26-sGP, a recombination reaction between the entry clone pDONR221-sGP and each individual destination vector was carried out using LR Clonase (Invitrogen). Briefly, 150 ng of the pDONR221-sGP plasmid was added to each of the three pDEST plasmids as separate experiments using a 10 μL LR Clonase reaction mixture. After incubation for 1 h at 25 °C, the *E. coli* TOP10 competent cells were transformed with 1 μL of each of the recombinant plasmids described above. Colonies were selected on a Luria–Bertani (LB) agar plate containing 100 μg/mL ampicillin. Positive clones were identified by PCR and confirmed by DNA sequencing. Single positive colonies from the transformation plate were then used to inoculate 5 mL of LB broth supplemented with ampicillin and incubated overnight at 37 °C with shaking, and then stored as glycerol stocks at −80 °C.

Bacterial expression and purification of recombinant sGP

Bacterial expression of recombinant sGP fusion protein: For bacterial expression *E. coli* Rosetta 2 Competent Cells (Novagen) were freshly transformed with the pDEST527-sGP expression vector. Small-scale expression (50 mL LB broth supplemented with 100 μg/mL ampicillin) was carried out at 37 °C with shaking at 250 rpm overnight. Next, 20 mL of the overnight culture were used to inoculate a 1 L culture and expression of recombinant sGP protein was induced by the addition of IPTG (Isopropyl β-D-1-thiogalactopyranoside) to 1 mM when the cell density OD_{600} reached 0.5. The cell culture was then allowed to grow at 37 °C for an additional 4 h post induction. Cells were then harvested by centrifugation at 4 °C for 10 min at 5000*g* and the cell pellets were frozen and stored at −80 °C. The solubility of recombinant sGP fusion proteins was analyzed using CelLytic B Plus Kit (Sigma-Aldrich) following manufacturer's directions.

Purification and refolding of bacterial expressed sGP fusion protein: To isolate inclusion bodies, 10 g of frozen cell pellets expressing 6×His-tagged full-length sGP were resuspended in 100 mL of PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM $Na_2HPO_4 \bullet 2H_2O$, and 1.76 mM KH_2PO_4), pH 7.4 containing 5 mM DTT (Dithiothreitol), 5 mM EDTA, 1 mM PMSF (phenylmethylsulfonyl fluoride), the Roche Complete Protease Inhibitor Cocktail, and 1 mg/mL lysozyme. The lysis suspensions were incubated for 30 min at room temperature with gentle agitation, followed by sonication on ice for 3 min. Cell lysates were cleared by centrifugation at 18,000 rpm (JA-25.50 Rotor in Beckman Coulter Avanti® J-E centrifuge) for 30 min at 4°C. The resulting pellets containing inclusion bodies were resuspended in 50 mL of wash buffer (50 mM Tris–HCl, pH 8.0, 2% Triton X-100 (v/v), 500 mM NaCl, 1 mM 2-mercaptoethanol, 1 mM PMSF, and 2 M urea) and centrifuged at 18,000 rpm for 20 min at 4°C. This step was repeated

three times with 50 mL of a Triton X-100 free wash buffer used for the final wash. Following the final centrifugation, the washed inclusion bodies were solubilized in 10 mL of solubilization buffer containing 6 M urea, 50 mM Tris–HCl, pH 8.0, and 1 mM 2-mercaptoethanol by incubation at room temperature for 6 h. The supernatant containing denatured sGP proteins were cleared from insoluble material by centrifugation at 18,000 rpm for 30 min at 4 °C, followed by filtration through a 0.45 μm syringe filter.

His-tagged sGP protein solubilized in 6 M urea was further purified with a 5 mL HisTrap™ HP Immobilized Metal Affinity Chromatography (IMAC) column connected in AKTA Prime system (Amersham–Pharmacia Biotech). The denatured protein samples were loaded onto the column pre-equilibrated with chromatography buffer A containing 6 M urea, 50 mM Tris–HCl pH 8.0, 1 mM 2-mercaptoethanol, and 20 mM imidazole. After sample loading, the column was washed with at least ten column volumes of buffer A to remove unwanted proteins. The His-tagged sGP fusion protein was then eluted using a 6-column volume linear gradient (20–500 mM) of imidazole in buffer A. Pooled elution fractions containing sGP protein were either stored at −80 °C or used directly for sGP refolding.

A stepwise dialysis procedure was used to refold the *E. coli*-expressed sGP. Briefly, 10 mL of purified denatured sGP was diluted with chromatography buffer A to approximately 1 mg/mL. The urea denaturant was then slowly removed by a series-dialysis using buffers of decreasing urea concentration 6 M→ 3 M→ 2 M→ 1 M→ 0.5 M→ 0 M urea in 20 mM Tris–HCl pH 8.4 containing 50 mM NaCl, 5 mM DTT, 5 mM EDTA, 5% glycerol, 5 mM reduced glutathione, and 0.5 mM oxidized glutathione. After the final overnight dialysis in urea free buffer, the sample was centrifuged at 18,000 rpm for 30 min at 4 °C to removed precipitated proteins. SDS PAGE (Coomassie blue staining) was used to determine the purity of refolded

His-sGP protein. Protein concentration was measured using the BCA Protein Assay Kit (Pierce) according to the manufacturer's instructions.

Baculovirus expression and purification of recombinant sGP

sGP baculovirus production for insect expression: The pDEST10-sGP plasmid described above was used to generate recombinant bacmids and baculovirus carrying the *sGP* gene. The bacmid was produced following the manufacturer's procedures described for the Bac-to-Bac system (Invitrogen). Briefly, the pDEST10-sGP plasmid was transformed into DH10Bac competent cells containing the baculovirus genome bacmid. After screening, white colonies containing recombinant bacmid containing the sGP insert were selected. The recombinant bacmids were purified and verified by PCR for the presence of the *sGP* gene.

Spodoptera frugiperda Sf9 cell maintenance: *Spodoptera frugiperda* (Sf9) cells adapted to HyClone SFX-Insect serum free medium (Thermo Scientific) supplemented with 1% Fetal bovine serum (FBS) and 1× gentamicin were cultured at 27 °C in either monolayer or suspension culture. Sf9 suspension cultures were maintained at a cell density of $1 – 6 \times 10^6$ cells/mL as 50 mL cultures in 250 mL shaker flasks under shaking at 95-120 rpm.

Insect cell transfection and recombinant virus isolation: To generate baculovirus for sGP intracellular production, newly attached monolayer Sf9 cells in 6-well plates (1×10^6 cells/well) were transfected with 1 μg of the PCR-verified recombinant sGP bacmid DNA using Cellfectin II™ (Invitrogen) according to the manufacturer's directions. Briefly, the Sf9 cells were incubated in 2 mL of the bacmid/Cellfectin transfection mixture diluted with HyClone SFX-Insect serum free medium for 5 h at 27 °C. After removal of the transfection mixture, 2 mL of

medium supplemented with 10% FBS and 1× gentamicin was added to each well. The medium containing the sGP baculovirus was collected 5 days post-transfection and reserved as the P1 viral stock. Next, a higher titer P2 viral stock was produced follows: 50 mL of Sf9 cells at a cell density of 2×10^6 cells/mL were infected with 1 mL of the P1 viral stock and the supernatant containing the P2 viral stock was harvested 3 days post-infection. The resulting P2 viral stock was then was for large-scale sGP expression.

Expression and purification of recombinant sGP in insect cells: For large scale sGP intracellular expression, 500 mL of the Sf9 suspension culture (2.5×10^6 cells/mL) flask maintained in supplemented SFX-Insect medium was placed in a 2 L shaker flask and infected with 500 µL of the P2 virus stock. The mixture was then cultured at 27 °C with shaking (95 rpm). Infected cells were harvested at 72 h post-infection and recovered from cultures by centrifugation at 2000*g* for 10 min. The cell pellets were then washed twice with ice-cold PBS and frozen stored at −80°C.

For purification, the cell pellets containing sGP fusion protein were resuspended in 1/40 the volume of original cell culture with ice-cold lysis buffer (PBS, pH 7.4, 5% v/v glycerol, 300 mM NaCl, 20 mM imidazole, and the Roche Complete Protease Inhibitor Cocktail). Crude cell extracts were prepared by a treatment consisting of 5 passes through a French press operated at 4 °C. Lysates were cleared by centrifugation at 18,000 rpm (JA-25.50 Rotor in Beckman Coulter Avanti® J-E centrifuge) for 30 min at 4°C and the supernatant was subjected to Nickel affinity chromatography, followed by ion exchange chromatography, HiTrap SP chromatography followed by HiTrap Q chromatography as outlined in Figure A.2. Briefly, the clarified, crude lysate was incubated with pre-equilibrated Ni Sepharose 6 Fast Flow resin (1 ml beads per 25 mL lysate) for 2 h at 4 °C. The Ni resin was then packed into an empty column and washed twice

with 20 resin bed volumes of Binding Buffer (10 mM phosphate buffer, pH 7.4, 5% v/v glycerol, 300 mM NaCl, 20 mM imidazole). The bound sGP protein was then eluted with 10 resin bed volumes of binding buffer containing 500 mM imidazole. Immediately following elution, the sGP eluate was pooled and dialyzed overnight against 10 mM phosphate buffer, pH 7.4, containing 5% v/v glycerol and 50 mM NaCl (chromatography Buffer B). The dialyzed sGP was further purified by ion exchange chromatography using a 5 mL HiTrap Q column equilibrated with chromatography Buffer B. The flow-through fractions containing the sGP protein from HiTrap Q column were then pooled and loaded onto a HiTrap SP column that had been pre-equilibrated with chromatography Buffer B. After sample loading, the column was washed with ten column volumes of Buffer B. The sGP fusion protein was eluted using 50 mL of chromatography Buffer B containing 1 M NaCl. The eluted fractions containing purified sGP were then aliquoted and stored frozen at −80 °C.

Mammalian cell expression and purification of recombinant sGP

Preparation of maxiprep plasmid DNA for mammalian expression: For plasmid preparation, a 500 mL LB broth culture was inoculated from a single colony or glycerol stock of pDEST26-sGP and incubated with shaking (220 rpm) for 16 h at 37 °C. Plasmid DNA was prepared using a QIAGEN Plasmid Plus Maxi Kit (Qiagen) according to the manufacturer's directions. The quantity and quality of plasmid DNA were verified by UV-spectrometry.

HEK 293T Cell culture and transient transfection: HEK 293T cells were grown in 150 mm cell culture dishes (BD Biosciences) in a 37 °C humidified atmosphere containing 5% CO_2, using Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal bovine

serum (HyClone). On the day of transfection, cells were approximately 80% confluent. The pDEST26-sGP plasmids were transfected into 293T cells using the QuickShuttle™ transfection reagent (KBQ-BIO) according to the manufacturer's instructions. Briefly, 30 µg of plasmid DNA and 50 µL transfection reagent were pre-diluted in 1 mL of unsupplemented DMEM respectively. The two solutions were mixed and incubated for 15 min at room temperature. The resulting 2 mL of the QuickShuttle™ transfection reagent - DNA mixture were added into the monolayer cell culture in a drop-wise manner and the cell cultures were allowed to grow for an additional 48 h. The transfected cells were gently washed twice with ice-cold PBS and frozen at -80 °C until use.

Purification of recombinant sGP from 293T cells: The purification protocol for insect-expressed sGP was used with only minor modification. Briefly, the cell lysates were centrifuged for 30 min at 18,000 rpm at 4°C. The resulting supernatant was incubated with Ni Sepharose 6 Fast Flow resin for 2 h at 4 °C as outlined above. The resin was then extensively washed and packed into an empty column. The sGP protein was then eluted as using 500 mM imidazole and further purified by ion exchange chromatography as described above.

SDS–PAGE and Western blot analysis

The sGP proteins produced in this study were analyzed by SDS–PAGE (4-20% gradient, BIO-RAD) and stained for protein with Coomassie blue. For Western blot analysis, protein samples were separated by SDS–PAGE and electrophoretically transferred to PVDF *(*Polyvinylidene fluoride*)* membranes (50 V, constant voltage) for 1 h using a Criterion™ Blotter (BIO-RAD). Membranes were then incubated in PBST blocking buffer (PBS, 0.2% Tween-20,

5% BSA) at room temperature for 1 h. Blocked membranes were probed with an anti-6×His antibody (His-Probe, Santa Cruz Biotechnology) diluted 1:5000 in blocking buffer for 1 h at room temperature. The membranes were then washed three times with PBST for 10 min each. The washed membranes were next incubated with a rabbit anti-mouse IgG-HRP (1:5000, Santa Cruz Biotechnology) for 30 min and washed again. Western Blotting using the Luminol Reagent (Santa Cruz Biotechnology) on x-ray film was used to detect the presence of recombinant sGP protein

RESULTS AND DISCUSSION

Cloning and vector construction

In this study, one of the most commonly used high-throughput cloning methods – Gateway® Technology (Invitrogen) was used to generate identical sGP vectors for each expression system. Gateway® Technology provides easy and rapid access to multiple expression options from a single entry clone of target gene, enabling access to virtually any expression system [33]. For this reason, PCR primers corresponding to the Ebola virus (Zaire) sGP open reading frame were designed to be compatible with Gateway® cloning. The constructs included a hexahistidine (6xHis) purification tag fused at the N-terminus of *sGP* gene product to facilitate protein purification by immobilized metal affinity chromatography. In addition, a TEV protease cleavage site was introduced to allow removal of 6xHis tag. Purified sGP PCR products were inserted into the Gateway® entry vector pDONR221 by BP reaction. The correct *sGP* gene sequence was confirmed by DNA sequencing. Since the sGP entry clone is available, we need

only to carry out LR reactions to subclone *sGP* gene into individual expression vector for *E. coli*, baculovirus /insect cell, and mammalian cell expression systems. By homologous recombination using LR Clonase, sGP cDNA was successfully cloned in each Gateway destination vector as verified by DNA sequencing. In this study, three Gateway destination vectors (pDEST527, pDEST10 and pDEST26) were chosen for each of the three expression systems (*E. coli*, baculovirus/insect cell, and mammalian cell) so that all the destination vectors would contain same N-terminal tags to ensure that all expressed sGP proteins have an identical amino acid sequence.

Expression and purification of sGP from *E. coli*

In order to produce recombinant sGP protein, the plasmid pDEST527-sGP was transformed and expressed in *E. coli* Rosetta 2 competent cells as shown in Figure A.3.a. Solubility analyses showed that sGP was insoluble and expressed as inclusion bodies. Several expression parameters, including post-induction temperature (18 or 37°C), time of induction (1-4h for expression at 37°C, overnight for expression at 18°C), the concentration of IPTG (0.1–1mM), and *E. coli* expression strains were evaluated in order to increase protein solubility. However, none of the conditions used showed a significant improvement in protein solubility (data not shown). The analyses of the insoluble fractions by SDS–PAGE showed recombinant sGP is expressed with molecular weight of approximately 43 kDa corresponding to the fusion of 6×His tag (~1 kDa) and full-length sGP (~42 kDa); The identity of expressed proteins was also confirmed by Western bolt using the monoclonal Anti-6×His antibody.

Inclusion bodies containing insoluble sGP in the cell lysate were pelleted by centrifugation and washed with urea and Triton X-100 to remove *E. coli* membrane and cell wall material. The urea concentration selected in the wash buffer was 2 M, because higher concentrations result in partial solubilization of the recombinant protein. We also included 2% Triton X-100 in the wash buffer to help extract lipid and membrane-associated proteins. A final wash step with buffer lacking Triton X-100 ensures the removal of excessive detergent from the inclusion bodies preparation. The denaturing agent, 6 M urea, a common denaturing agents for insoluble proteins, was used to solubilize sGP in this study. Higher concentrations of urea were not needed since 6 M urea appeared to be very efficient in solubilizing sGP from the inclusion bodies.

Results from immobilized metal affinity chromatography as reflected in the chromatogram and SDS-PAGE analysis are shown Figure A.3.b. indicate that sGP could be effectively purified by Ni affinity chromatography in the presence of 6 M urea. The single peak (blue line) observed in Figure A.3.b at ~1200 mAU (OD_{280}) showed that sGP was eluted with approximately 250 mM imidazole in the gradient elution profile. This purification step produced ~10 mg denatured sGP for every gram of starting bacterial cells. The SDS–PAGE analysis (Figure A.3.a) showed that the sGP could be effectively purified by Ni affinity chromatography in the presence of 6 M urea, and the additional protein bands with lower MW may be an indicative of sGP degradation.

Since the sGP from inclusion bodies was purified in the presence of 6 M urea as a denatured protein, it must be refolded for further structural or immunologic studies. The refolding was accomplished by stepwise dialysis with decreasing urea concentration (6 M to 0

M). The protein that precipitated during dialysis was removed by centrifugation as soon as the protein solution became "cloudy" or the precipitates became visible. In brief, 20 mL of denatured sGP (~ 1 mg/mL) was dialyzed in steps against 2 L of dialysis buffer containing decreasing amounts of urea. Each step of the serial dialysis was carried out for a minimum of 8 h at 4°C. The overall yield of refolded sGP protein was approximately 10 mg/L of bacterial culture (Figure A.3.c).

Expression and purification sGP from baculovirus/insect system

The insect cell is an excellent eukaryotic expression system with many advantages that bacterial systems cannot offer, such as producing functional recombinant proteins and correct post-translational modifications. The *sGP* gene was cloned into the pDEST10 vector and transformed into *E.coli* DH10Bac strain to generate sGP bacmids. The correct bacmids carrying sGP gene was then transfected into Sf9 insect cells using Cellfectin II as the transfection reagent. The baculovirus was then allowed to amplify in the initial insect cell culture for 5 days and recovered in culture medium. Technically, the first-generation P1 baculovirus is of low titer and used to infect fresh insect cells to produce higher titer virus stock (~ 10^9 pfu/ml) after 1-2 generations. The expression strategy for sGP protein in this study was based on the results of a pilot expression test, in which a 50 mL cell suspension cultured in the supplemented SFX-Insect medium at 27°C with cells infected at varying MOIs (multiplicity of infection) of 1-10. The expression was tested at different times post-infection (24h, 48h and 72h) by SDS-PAGE and Western blot. Optimal expression conditions were determined for the following scale-up expression that the P2 virus with an estimated titer of 0.5 – 1 × 10^9 plaque forming units (pfu)/mL was used to infect Sf9 cultures at the MOI of 1 with the cell cultures harvested 72 h

post-infection. Also in this study we expressed sGP cytoplasmically to avoid post-translational glycosylation of sGP in insect cells. To purify cytoplasmically expressed sGP, the cell cultures were pelleted by centrifugation for 10 minutes at 1000xg. The resulting cell pellets were then lysed by the French press instead of using lysis buffers containing high concentration of detergents such as Triton X-100 and NP-40 to avoid difficulties in further immunological or crystallization studies. The cell lysates were cleared by centrifugation at 18,000 rpm for 30 minutes. Purification of insect cell expressed sGP was carried out using the chromatographic procedures including Ni affinity resins and two types of ion exchange chromatography columns. After purification, the sGP exhibited a single major band by SDS-PAGE as shown in Figure A.4 with a molecular weight of ~43 kDa and suggesting that the protein is not glycosylated as expected. Typically, this expression/purification protocol yielded approximately 1 mg of sGP per liter of insect cell culture.

Expression and purification of sGP from HEK 293T cells

For mammalian expression, we cloned *sGP* gene into pDEST26 vector. The expression vector pDEST26-sGP was transfected into 293T cells using transfection reagent. Forty-eight hours post-transfection, cells were harvested and lysed. Transient expression of sGP in 293T cells was confirmed by SDS–PAGE and Western blot analyses. Figure A.5.a shows the presence of a major band of ~43 kDa in the transfected cell extract that was not observed in the control (un-transfected) cell extract suggesting that this band represents the recombinant sGP protein. We also tried to express sGP in Chinese hamster ovary (CHO) cells, but CHO cells failed to express the protein, as indicated in Figure A.5.b. To purify 293T cell expressed sGP the same purification protocol used for insect cell expressed protein was employed. As indicated in Figure

A.5.c, sGP were purified successfully with a molecular weight of ~43 kDa without any sign of glycosylation. The purity of these recombinant proteins was estimated to be approximately 85%.

CONCLUSION

The overall goal of this work is to combine the use of biochemistry, biophysics, structural biology, and immunology to investigate the immunogenicity of recombinant vaccine proteins generated *in vitro* using prokaryotic and eukaryotic expression systems. As a general rule in vaccine development, bacterial proteins should preferably be made in transfected bacterial cells, and human viral antigens, *i.e.* glycoproteins, in mammalian cells, because of the substantial differences in properties, such as post-translational modifications in different cell types. However, the roles that these differences play in the structure and/or immunogenicity of a recombinant protein are only now beginning to be understood. In addition, there is an emerging role of using structural biology to aid in vaccine development.

The work presented here describes the expression and purification of Ebola sGP in three different expression systems (bacterial *E. coli,* baculovirus insect cell and mammalian expression systems). The sGP protein produced in *E. coli* was in a really high yield as an insoluble 43 kDa polypeptide. After purification in the presence of 6 M urea, the protein was refolded and will be used to perform subsequent immunologic and structural studies. In contrast, recombinant sGP proteins expressed intracellularly in insect cells and mammalian cells were found to be non-glycosylated soluble. The intracellular expression strategy used in this study was to facilitate the subsequent immunologic and structural studies since extensive sGP glycosylation may interfere

with these studies. However, to test our hypothesis that source of the recombinant protein used to produce the vaccine would affect the protein structure and subsequent antigen-specific immune response, the examination of the immune responses of the host animals will be followed immediately using sGP produced in each expression system. Furthermore, the results from this study and subsequent examination of the immune responses may be a valuable tool to study the structure–immunogenicity relationship for vaccine development.

Figure A.1. The Ebola virus genome and expression in host cells. Adapted from [5].

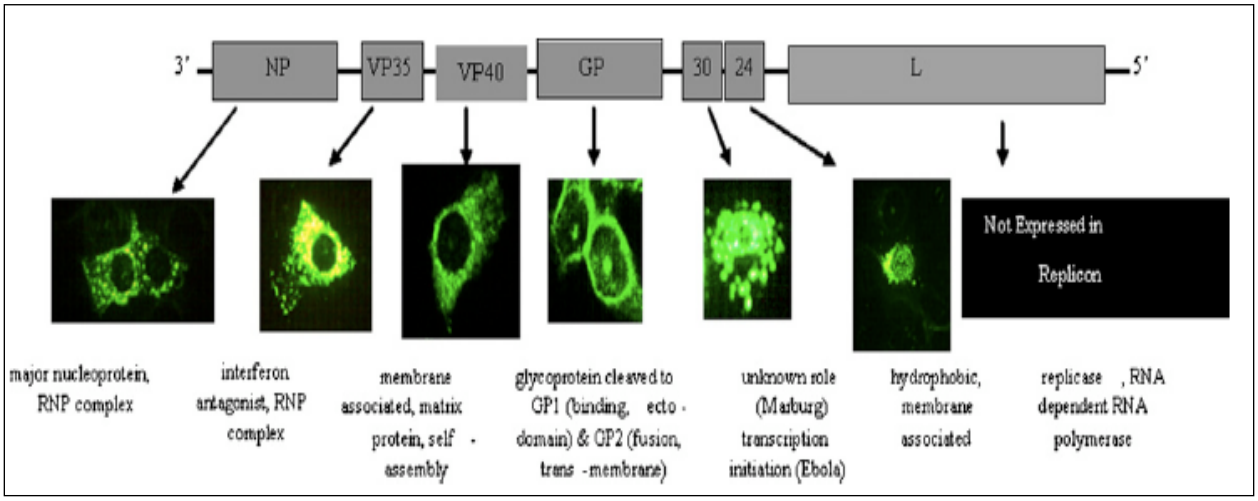


Table A.1. A summary of the sGP expression systems in this study.

Expression systems	Cells	Method	Gateway vectors	Gene
E.coli	Rosetta 2	Plasmid/induction	pDEST527	Wild type
Insect	Sf9	Baculovirus/infection	pDEST10	Wild type
Mammalian	HEK 293T	Plasmid/transfection	pDEST26	Wild type

Figure A.2. Flow chart of expression and purification of recombinant sGP from insect Sf9 cells.

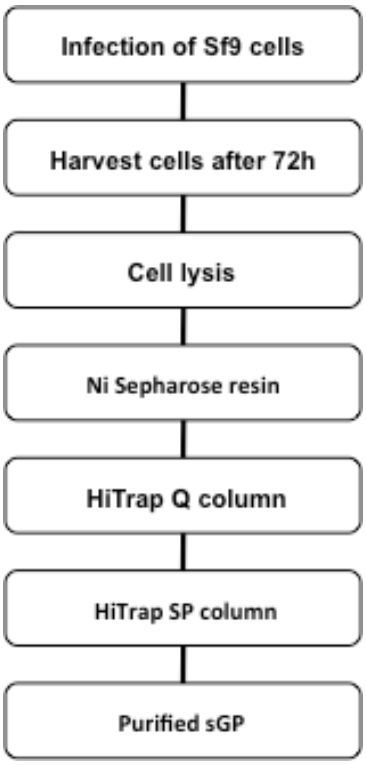
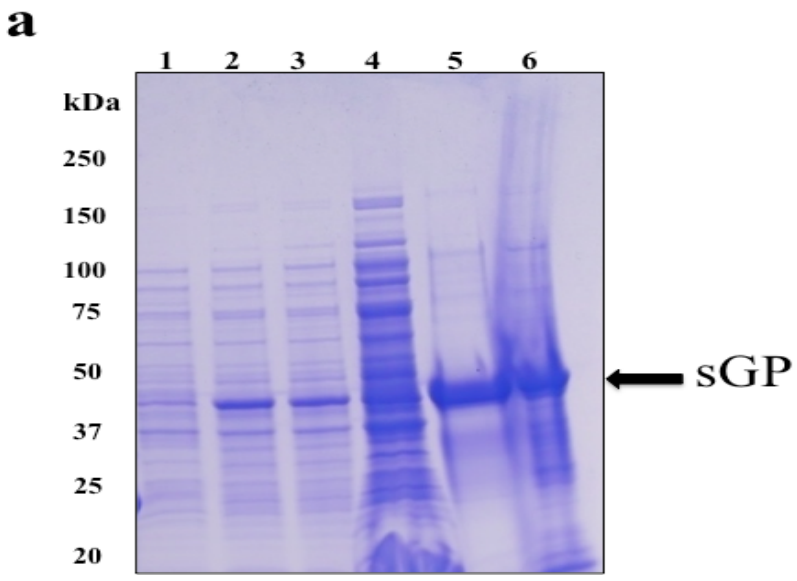


Figure A.3. Purification and refolding of recombinant sGP produced by bacterial expression.

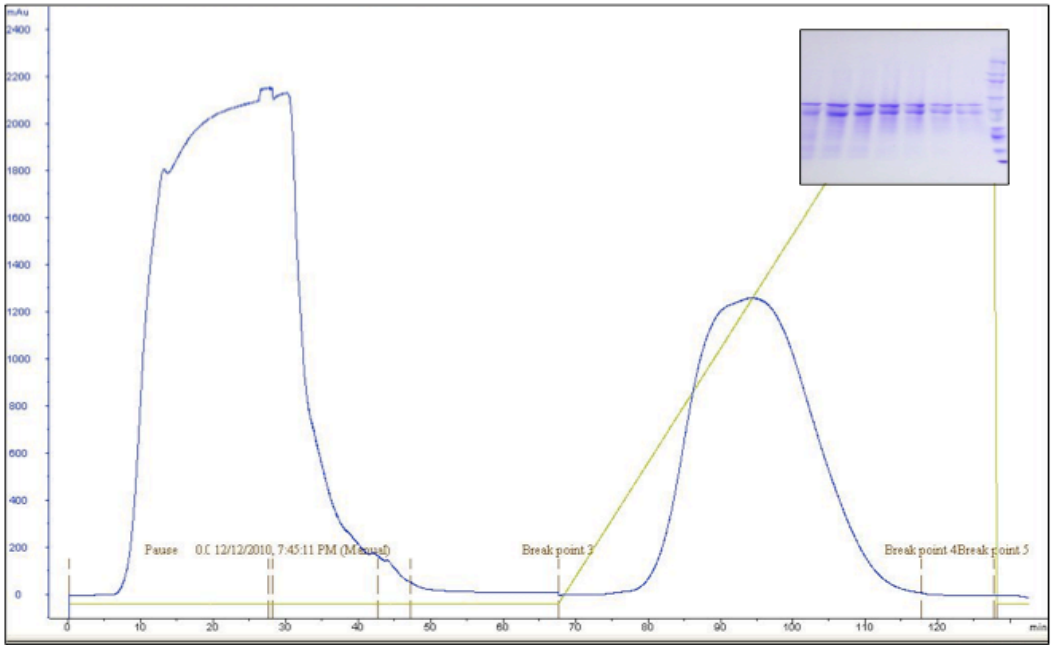
a, SDS–PAGE analysis of expression and solubility test of sGP in *E.coli*, uninduced cells (lane 1) and IPTG induced cells (lane 2-3) showing expression of sGP with the MW of ~43 kDa. Lane 4-6: Supernatant of cell lysate, pellet of cell lysate dissolved in 6 M urea, and pellet of cell lysate dissolved in Laemmli sample buffer; sGP expressed in *E. coli* is insoluble and present in the inclusion bodies. The bacteria expressed sGP is indicated by the arrow.

b, A chromatogram of HisTrap column purification of His-tagged sGP in the presence of 6 M urea (blue line). The His-tagged sGP was eluted by an imidazole gradient (20 – 500 mM, green line); insert shows the SDS-PAGE analysis of elution fractions containing purified sGP.

c, SDS-PAGE analysis of refolded sGP, as indicated by the arrow.



b


mAu
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Pause 0.(12/12/2010, 7:45:11 PM (Manual)
Break point 3
Break point 4Break point 5
0
10
20
30
40
50
60
70
80
90
100
110
120
min

c

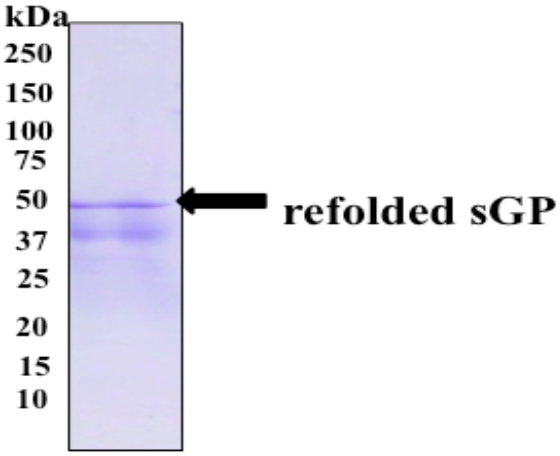

kDa
250
150
100
75
50
37
25
20
15
10
refolded sGP

Figure A.4. Expression and purification recombinant sGP expressed in insect/baculovirus system.

a, A western blot analysis of expression of recombinant sGP in Sf9 insect cells using anti-6×His mAb, showing the expression of sGP with the MW of ~43 kDa.

b, SDS-PAGE analysis of recombinant sGP purified by final HiTrap SP column chromatography (as described in Experimental methods). Lane 1 and 2: the flow-through fractions of final HiTrap SP column chromatography；lane 3-5: the HiTrap SP column chromatography elution fractions containing purified sGP， as indicated by the arrow.

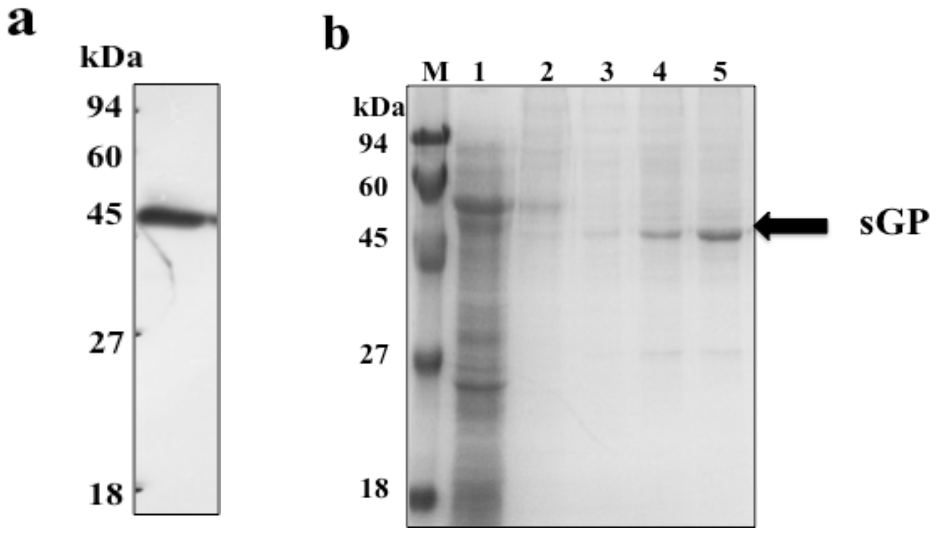
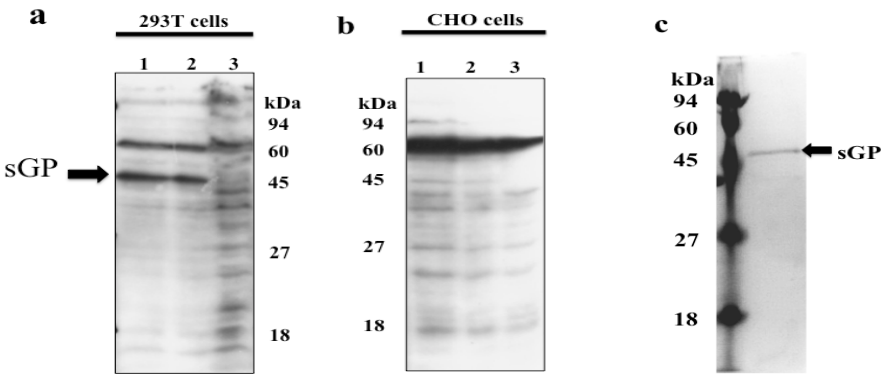


Figure A.5. Expression and purification recombinant sGP in mammalian HEK 293T cells.

a, A western blot analysis of expression of recombinant sGP in 293T cells using anti-6×His mAb, showing the expression of sGP with the MW of ~43 kDa. Lane 1-2: pDEST26-sGP transfected 293T cells at 48 h post-transfection. Lane 3: control 293T cells without transfection.

b, A western blot analysis of expression of recombinant sGP in CHO cells using anti-6×His mAb, showing that the sGP was not expressed in CHO cells. Lane 1: control CHO cells without transfection; lane 2-3: pDEST26-sGP transfected CHO cells at 48 h post-transfection.

c, SDS-PAGE analysis of purified sGP, as indicated by the arrow.



LITERATURE CITED

1. Bray, M., Defense against filoviruses used as biological weapons. Antiviral Res, 2003. 57(1-2): p. 53-60.

2. Geisbert, T.W. and P.B. Jahrling, Towards a vaccine against Ebola virus. Expert Rev Vaccines, 2003. 2(6): p. 777-89.

3. Hart, M.K., Vaccine research efforts for filoviruses. Int J Parasitol, 2003. 33(5-6): p. 583-95.

4. Feldmann, H., et al., Ebola virus: from discovery to vaccine. Nat Rev Immunol, 2003. 3(8): p. 677-85.

5. Reed, D.S. and M. Mohamadzadeh, Status and challenges of filovirus vaccines. Vaccine, 2007. 25(11): p. 1923-34.

6. Rotz, L.D., et al., Public health assessment of potential biological terrorism agents. Emerg Infect Dis, 2002. 8(2): p. 225-30.

7. Sullivan, N., Z.Y. Yang, and G.J. Nabel, Ebola virus pathogenesis: implications for vaccines and therapies. J Virol, 2003. 77(18): p. 9733-7.

8. Stroher, U., et al., Infection and activation of monocytes by Marburg and Ebola viruses. J Virol, 2001. 75(22): p. 11025-33.

9. Wahl-Jensen, V.M., et al., Effects of Ebola virus glycoproteins on endothelial cell activation and barrier function. J Virol, 2005. 79(16): p. 10442-50.

10. Bausch, D.G., et al., Treatment of Marburg and Ebola hemorrhagic fevers: a strategy for testing new drugs and vaccines under outbreak conditions. Antiviral Res, 2008. 78(1): p. 150-61.

11. Feldmann, H., et al., Filovirus-induced endothelial leakage triggered by infected monocytes/macrophages. J Virol, 1996. 70(4): p. 2208-14.

12. Wahl-Jensen, V., et al., Role of Ebola virus secreted glycoproteins and virus-like particles in activation of human macrophages. J Virol, 2005. 79(4): p. 2413-9.

13. Hoenen, T., et al., Ebola virus: unravelling pathogenesis to combat a deadly disease. Trends Mol Med, 2006. 12(5): p. 206-15.

14. Vastag, B., Ebola vaccines tested in humans, monkeys. JAMA, 2004. 291(5): p. 549-50.

15. Hensley, L.E., et al., Demonstration of cross-protective vaccine immunity against an emerging pathogenic Ebolavirus Species. PLoS Pathog, 2010. 6(5): p. e1000904.

16. Lee, J.E., et al., Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature, 2008. 454(7201): p. 177-82.

17. Lee, J.E., et al., Complex of a protective antibody with its Ebola virus GP peptide epitope: unusual features of a V lambda x light chain. J Mol Biol, 2008. 375(1): p. 202-16.

18. Lee, J.E. and E.O. Saphire, Neutralizing ebolavirus: structural insights into the envelope glycoprotein and antibodies targeted against it. Curr Opin Struct Biol, 2009. 19(4): p. 408-17.

19. Mellquist-Riemenschneider, J.L., et al., Comparison of the protective efficacy of DNA and baculovirus-derived protein vaccines for EBOLA virus in guinea pigs. Virus Res, 2003. 92(2): p. 187-93.

20. Sanchez, A., et al., The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proc Natl Acad Sci U S A, 1996. 93(8): p. 3602-7.

21. Takada, A., et al., A system for functional analysis of Ebola virus glycoprotein. Proc Natl Acad Sci U S A, 1997. 94(26): p. 14764-9.

22. Feldmann, H., et al., Biosynthesis and role of filoviral glycoproteins. J Gen Virol, 2001. 82(Pt 12): p. 2839-48.

23. Sanchez, A., et al., Biochemical analysis of the secreted and virion glycoproteins of Ebola virus. J Virol, 1998. 72(8): p. 6442-7.

24. Yang, Z., et al., Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. Science, 1998. 279(5353): p. 1034-7.

25. Falzarano, D., et al., Ebola sGP--the first viral glycoprotein shown to be C-mannosylated. Virology, 2007. 368(1): p. 83-90.

26. Falzarano, D., et al., Structure-function analysis of the soluble glycoprotein, sGP, of Ebola virus. Chembiochem, 2006. 7(10): p. 1605-11.

27. Ritchie, G., et al., Identification of N-glycans from Ebola virus glycoproteins by matrix-assisted laser desorption/ionisation time-of-flight and negative ion electrospray tandem mass spectrometry. Rapid Commun Mass Spectrom, 2010. 24(5): p. 571-85.

28. Feldmann, H., et al., Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. Virology, 1994. 199(2): p. 469-73.

29. Dowling, W., et al., Influences of glycosylation on antigenicity, immunogenicity, and protective efficacy of ebola virus GP DNA vaccines. J Virol, 2007. 81(4): p. 1821-37.

30. Wei, X., et al., Antibody neutralization and escape by HIV-1. Nature, 2003. 422(6929): p. 307-12.

31. Koch, M., et al., Structure-based, targeted deglycosylation of HIV-1 gp120 and effects on neutralization sensitivity and antibody recognition. Virology, 2003. 313(2): p. 387-400.

32. Hevey, M., et al., Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants. Virology, 1997. 239(1): p. 206-16.

33. Hartley, J.L., G.F. Temple, and M.A. Brasch, DNA cloning using in vitro site-specific recombination. Genome Res, 2000. 10(11): p. 1788-95.