Role of Reverse Transcriptase and APOBEC3G in Survival of Human Immune Deficiency Virus -1 Genome

Raj Kumar Soni*, Amol Kanampalliwar and Archana Tiwari
School of Biotechnology, Rajiv Gandhi Proudyogiki Vishwavidyalaya (State Technological University), Madhya Pradesh, India

Abstract
Development of an effective vaccine against HIV-1 is a major challenge for scientists at present. Rapid mutation and replication of the virus in patients contribute to the evolution of the virus, which makes it unquerable. Hence a deep understanding of critical elements related to HIV-1 is necessary. Errors introduced during DNA synthesis by reverse transcriptase are the primary source of genetic variation within retroviral populations. Numerous current studies have shown that apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) proteins mediated sublethal mutagenesis of HIV-1 proviral DNA contributes in viral fitness by accelerating human immunodeficiency virus-1 evolution. This results in the loss of the immunity and development of resistance against anti-viral drugs. This review focuses on the latest biological, biochemical, and structural studies in an attempt to discuss current ideas related to mutations initiated by reverse transcriptase and APOBEC3G. It also describes their effect on immunological diversity and retroviral restriction, and their overall effect on the viral genome respectively. A new procedure for eradication of HIV-1 has also been proposed based on the previous studies and proven facts.

Keywords: HIV-1; Reverse transcriptase; Recombination; APOBEC3G; Viral infectivity factor; Mutation

Introduction
AIDS was first described clinically in 1981 in the United States. Since then more than 34 million people are diagnosed with AIDS. After three decades of research, an effective vaccine for HIV-1 is yet to design. Modern techniques have delievered us with antiretroviral therapies, which can mollify AIDS and increase the life expectancy of patients. These drug treatments have many side effects once introduced they must be continued throughout the life of the patient. Drug toxicity and the development of resistance for ongoing therapy make them unbearable. High mutation rate and ability of HIV-1virions to hide inside the body of the host contribute in development of drug resistance. Reverse transcriptase (RT) plays an important role in replication of viral genome. However, the lack of any proofreading mechanism makes this process highly error prone and most significant for viral diversification. Majority of the mutations in the viral genome occurs at this step. APOBEC3G is a newly discovered enzyme with potential to inhibit transmission of the HIV. During reverse transcription, APOBEC3G enzyme can transform deoxycytidine (dC) into deoxouridine (dU) on the minus strand of viral DNA. This leads to the introduction of various errors at several unspecific positions of viral genome thus affecting a variety of biological activities in HIV-1 [1,2]. This review focuses on different mechanisms by which both these enzymes RT and APOBEC3G enhance the adaptability and evolution of the viral genome [3,4].

Reverse Transcriptase-Catalyst of the Reverse Transcription Reaction
RT is an asymmetric heterodimer made up of p66 and p51 subunits of different lengths and arrangements [5,6]. Study of diversity of HIV-1 suggests that genetic variations in four major groups and various subgroups in separate geographical distributions are the result of divergent evolution [7,8]. This spread was observed to be inconsistent based on the limited Group O epidemic in comparison to the Group M and Group N. Since the divergent group O evolution was not found to alter the structures or RT activity in comparison to that of HIV-1 group M subtype B hence it was used for classifying the majority of SIVs into two major groups or clusters. Based on the presence of Phenylalanine, Valine or Isoleucine at position 181 of the HIV-1, first group of RT was defined. Another group of RT included all HIV-1 isolates (group’s M, N, O, and P) as well as SIV sequences from gorillas (SIV-1) [9].

The “resistance” to Non-nucleoside reverse-transcriptase inhibitors (NNRTI) in group O viruses is intrinsic and arose during the divergent evolution of HIV-1 group’s M and O viruses in the simian/human immunodeficiency virus lineage of lentiviruses [10-13]. Intrinsic resistance in this lineage suggests a flexible accommodation of the NNRTI binding pocket of RT to genetic changes. Evidence for such flexibility might be best described by the relatively low fitness cost of NNRTI-resistant mutations K103N and Y181C compared to the higher fitness cost of other most drug-resistant mutations in HIV-1 group M subtype B isolates [14]. Drug resistance mutation (such as Y181C, A98G, K103N, V179E) in these regions which may result in an NNRTI-sensitive virus mostly occur with a cost of replicative fitness in both Group O and Group M as these are the same sites which distinguish lineages of group O and group M as well as other primate lentivirus species from each other.

Mode of Action of Reverse Transcription in HIV-1
The process of reverse transcription in HIV-1 retrovirus begins with the synthesis of the minus strand at the 5’ end of the plus strand of viral genome. At this position, 18 nucleotide long site, known as the protein binding site is present. At 3’end of human tRNALys3, 18 nucleotides long segment is situated, which is complementary to protein binding site and is essential for protein binding site. HIV-1 uses tRNALys3 as a primer for minus strand synthesis. Hybridization of protein binding site and tRNALys3 begins the reverse transcription process with the help of DNA polymerase function of RT using viral RNA as a template. As the DNA synthesis reaches 5’end of...
single stranded RNA, it forms first minus strand strong-stop DNA. Now RNase H function of RT comes into action and selectively breaks the 5' end of the viral RNA exposing the minus strand of recently synthesized single stranded package which carries two copies of the viral RNA genome[13-16]. One of the two single stranded RNAs hybridized with the newly synthesized single stranded DNA at 3' end through a strand transfer reaction known as minus-strand transfer. Subsequently, as the minus-strand synthesis continues along the length of the viral genome, so does RNase H degrades [15,16].

There are two distinct RNase H resistant sites one near the 3’ end of the RNA and other near the central region of the HIV-1 genome. These two specific purine-rich sequences (cPPT, 3PPT) are acknowledged as poly purine tract (PPT) and primers plus-strand DNA synthesis initiated from the 3’end of PPT. As minus-strand DNA synthesis nears the 5’end RT duplicates the first 18 nucleotides of the Lys3 tRNA primer setting the platform for plus-strand transfer. DNA synthesis stops when it encounters a modified A that RT cannot copy which leads to reverse-transcription of a portion of the primer tRNA consequent coding a DNA called plus-strand strong-stop DNA. Latter RNase H acts upon the tRNA-DNA hybrid and cleaves the RNA segment. This exposes primer binding site in sequence at plus-strand of DNA allowing it to anneal to the complementary sequences at or near the 3’end of minus-strand DNA thereby initiating the second strand transfer or second jump. DNA synthesis extends the minus and plus strands in opposite directions. Each strand of DNA serves as a template for the other strand producing a linear double stranded DNA with long terminal repeats (LTRs) at both ends [17-19] whole process is discussed in Figure 1.

Variations due to Reverse Transcriptase and RNA polymerase II

RNA polymerase II (RNA pol II) lacks a proofreading function which plays an important role in diversification of HIV-1 genome by causing polymerization errors during RNA pol II mediated DNA synthesis stops when it encounters a modified A that RT cannot

![Figure 1: HIV-1 Genome Reverse Transcription Process](image-url)

**Figure 1:** HIV-1 Genome Reverse Transcription Process. Step1: minus strand synthesis – 5’ end of the plus-strand RNA genome act as a primer for minus strand DNA synthesis. Step2: First strand transfer – minus strand strong stop DNA is formed due to RNaseH hydrolysis of RNA strand. Step3: minus strand synthesis - hybridization of minus strand DNA with R region at the 3’end of the ssRNA genome leads to further elongation of minus strand DNA. Step4: plus and minus strand synthesis – RNA and DNA strands keep on elongating after reaching to specific length RNaseH cleaves the RNA strand at certain regions. Step5: plus strand synthesis – using minus strand DNA minus strand synthesis begins. Step6: second strand transfer – polypurine tract (ppt) region is hydrolyzed by RNaseH releasing primer binding strand (PBS) of the plus strand DNA. Step7: strand displacement synthesis – PBS strand on the minus strand anneals with PBS present on plus strand DNA, which helps in continuing DNA synthesis. Step8: product formation – a linear double strand DNA is formed as a result of successive ligation and repair process, which contain long terminal repeats (LTRs) at both the ends.
transcription. RT and RNA pol II have a combined mutation rate of approximately $2 \times 10^{-5}$ per nucleotide per replication cycle [3,20-24]. Until now, separate contribution of RT and RNA pol II to the overall viral mutation rate has not been completely determined due to lack of any reliable technique to identify errors made by pol II and hence contributions of RNA pol II to the retroviral mutation rate is unknown. RT based errors are proposed to arise more frequently at specific positions known as hotspots.

Most of the *in vitro* experiments performed in various labs for identification of these hotspots either produced no patterns or in some cases generates different types of patterns of hotspots [21,24]. Possible explanations behind getting such difference in results are:

- Contribution of different types of viral and host proteins
- Application of distinct assay conditions and difference in purified RTs used during the assays.

Possible contribution of RNA pol II behind some of the hotspots (RT based mutation) [25]

According to a research performed by Hemelaar et al., about 20% of the currently circulating HIV-1 variants are inter-subtype recombinants proposing recombination as a major element in the evolution of the HIV-1 population [26].

**Variations due to Recombination**

Due to its ability to switch between two co-packaged RNAs, RT can generate chimeric DNA containing distinct portions of sequences from each of these two genomic RNAs. Only virions packaged with two genetically different RNAs are capable of generating a recombinant. Steps involved in the recombination process have been shown in Figure 2 in brief. There is not much information available related to double infection currently even though recent researches suggest that cells with more receptors are more susceptible to double infection by HIV-1 and occurs more frequently than expected when HIV-1 is transmitted via cell-mediated events [27,28].

Studies performed on HIV-1 recombination have suggested that the rate of recombination increases uniformly with the distances that separate the two alleles and attains its maximum possible value when both the alleles are separated by 1.3 kb [29]. Recombination rate is also affected by homology of sequence as recombination has more probability to occur when two co-packaged HIV-1 RNAs are from same subtype when compared with two co-packaged HIV-1 RNAs from different subtypes [30].

![Figure 2: Steps involved in recombination.](image-url)
Recombination process by shuffling and repeating existing mutations helps in increasing genetic diversity of HIV-1 viral population and production of variants, which are capable of surviving in any environment and evade host’s immune system [31]. Multidrug-resistant variants are also produced due to blending of drug-resistance mutations by the process of recombination [32,33]. Thus, recombination can affect HIV diversity within an individual and even globally and plays a critical role in placing hurdles in the development of effective vaccines and antiviral treatments. Recombination in HIV-1 occurs by switching mechanism, which has been discussed in Figure 3.

**APOBEC3G and its Significance in Cytidine Deamination Reaction**

APOBEC3G was first isolated during a study carried out by using a subtractive cDNA cloning approach in order to understand the Vif gene expression (HIV-1ΔVif) and its effects [34]. In humans, it includes eleven members to name a few of these are Activation-induced cytidine deaminase (AID), APOBEC1, APOBEC2, APOBEC3–H and APOBEC4 [35,36]. APOBEC3G and A3F proteins are extensively studied and recognized by researchers due to their ability to restrain retroviral infections [37]. APOBEC3G and A3F consists of two homologous catalytic domains, the N-terminal (CD1) and C-terminal (CD2) domains, each of these domains consist of a Zn$^{2+}$ coordination motif unique to cytidine deaminases (CTDAs) [38]. Conversely only one catalytic domain is present in A3H protein, and it consist of several haplotypes, out of these haplotypes II, V, and VII are reported to be highly active against HIV-1 [39-44].

During HIV-1 replication, APOBECs get entrenched into viral genome and modify minus-strand DNA by converting some of the cytosine present in the minus strand to Uracil, before the synthesis of plus strand and degradation of RNA. Analysis of several generations of HIV-1 genome sequences have revealed that presence of Vif (viral infectivity factor) gene in minus-strand of the reverse transcript of HIV-1 virions can affect the rate of deamination of cytosine’s (conversion of a cytidine into a uridine). Since minus strand serves as a template for the plus strand synthesis hence A is synthesized on the plus strand opposite to U instead of G causing G$\rightarrow$A mutation (Figure 4) [2,45-48]. At “hot-spots” of retroviral DNA deamination results in G$\rightarrow$A hypermutations. Such mutations are capable of terminating the coding and replicative ability of the virus [49,50].

In comparison to wild-type APOBEC3Gmediated, G$\rightarrow$A hypermutation could lead to 1,000-fold decline in infectivity of HIV-1 Virions lacking Vif [47,51-55]. Analytical studies have further reported that G$\rightarrow$A mutations induced by APOBEC3G proceeds in 5’$\rightarrow$3’ direction at a regular rate over the length of the genome [56]. Most research done on APOBEC3G and cytidine deamination reaction claims that cytosine deamination can inhibit HIV-1 replication in two possible ways:

1. It can degrade most reverse transcripts before their integration [47]. This degradation is expected to be triggered by DNA repair enzymes like uracil DNA glycosylase (UDG) and AP endonuclease.

2. G$\rightarrow$A mutations can generate several premature termination codons such as TAG, TGA and TAA, into the open reading frame of viral genome there by inhibiting next round of infection [56].

Many studies performed to analyze preferred target sequence of human APOBEC3G in both in vivo and in vitro assays suggest that human APOBEC3G more repeatedly target CC and less frequently targets TC, but rarely targets AC or GC [2,48,56,57]. This concludes that APOBEC3G mediated G$\rightarrow$A mutation mostly occurs in GG and less regularly in GA but never in GC and GT. In perspective of tetranucleotide CCCA/G was the most preferred target sequence by human APOBEC3G [56]. Analytical study of samples of HIV-1 patients suggests that GG and GA are favorite target sites of APOBEC3G for G$\rightarrow$A hypermutation [58-60].

**Vif Custodian against APOBEC3G**

Vif is very crucial protein and highly conserved in all lentiviruses with the exception of Equine Infectious Anemia Virus [61,62]. Currently, our knowledge related to structure of HIV-1 Vif protein is diminutive due to difficulty in expressing high levels of soluble recombinant proteins. Several considerably conserved motifs have been observed on alignment of Vif sequences of HIV-1, HIV-2 and SIV subtypes such as a tryptophan-rich stretch, a conserved HCCH domain and an SOCS-box [63]. According to a three-dimensional structural model of HIV-1, Vif based on APOBEC3G/Vif interaction, latter (Vif) consists of two functionally significant domains. N-terminal domain...

Figure 4: APOBEC3G Mediated Cytidine Deamination.

**Contribution of Vif and APOBEC3G Interaction in HIV-1 Variation**

Vif plays a crucial role in inhibition of APOBEC3G mediated cytidine deamination reaction as it can induce rapid degradation of APOBEC3G finagling which leads to its effective elimination from virions. Earlier researches have reported that proteasome inhibitors are capable of preventing Vif-mediated APOBEC3G degradation and could restore back APOBEC3G. These inhibitors can further direct encapsulation of active Vif-APOBEC3G complexes into the viral genome thus sustaining cytidine deaminase activity of APOBEC3G [52-54,71-73].

Yu et al. were first to describe the mechanism of Vif-induced APOBEC3G degradation with the help of SOCS box E3 ligase complex [73]. This complex modifies a large number of cellular proteins by attaching ubiquitin molecules such as signaling proteins, cell cycle proteins and cytokine receptors. The SOCS box protein helps in binding of E3 ligase to a substrate cellular protein and leads to ubiquitination of substrate by E2-ubiquitin and E3 ligase conjugating enzyme [74]. The SLQXLXLA motif in Vif is very important for ubiquitination process as was found to be similar to the central region of some SOCS boxes and mutation in this motif inhibits the Vif-induced APOBEC3G degradation [73]. Studies based upon mechanism of Vif-induced
APOBEC3G degradation have suggested that Vif is not capable of completely shielding the virus from effects of APOBEC3G and A3F leaving minor quantity of these enzymes encapsidated in the viral genome thus results in continuous low level of G→A mutations in viruses replicating in vivo [47,52,54,55,75].

APOBEC3G-mediated confinement of HIV-1 could lead to production of viable viral progeny with sub-lethal levels of G→A mutations. From recent studies, it has also been suggested that these types of mutations may contribute significantly to HIV-1 evolution, immune escape, drug resistance and pathogenesis [50]. It was proposed in earlier researches that because of G→A mutation caused by APOBEC3G in HIV-1 isolates, and infected patient’s defective Vif alleles are readily present and not all of these perceived defects can lead to a complete inhibition of function carried out by Vif therefore, some mutants are able to retain selectively neutralizing action against either APOBEC3G or A3F but not against both at the same time [76].

Detection of high adenosine content in lentiviral genomes and reports of detection of hyper-mutated genomes in vivo also advocates the role of cytidine deamination as an important source of viral diversification [59,77,78]. Hence from the above data we can recommend that APOBEC3G mediated cytidine deamination activity, which is well established as the host defense mechanism against HIV-1 virions that might have been altered to assist for evading immunological and pharmacological inhibition.

APOBEC3G as Sentinel or a Mole

Due to the irregular amount of G→A and effects of polymorphisms in Vif gene, hypermutation induced by APOBEC3G activity is not sufficient to limit HIV-1 infection, and sometimes it can actually increase HIV-1 diversification [79-83]. Recent studies also suggest that APOBEC3G is in fact, promoting HIV-1 diversification as a side effect to its failure in neutralizing viral genomes [50,84]. Out of all the drug resistance mutation sites discovered in HIV-1 many reside in APOBEC3G hotspots [85]. M184I mutation of RT is one such example, which resides in HIV-1 hotspot between TCCAT to TCUAT, introduced by APOBEC3G during HIV replication [86]. M184I mutation was reported to cause resistance to 2′, 3′-dideoxy-3′-thiacytidine (3TC) and to a lesser extent abacavir (ABC) and 2′, 3′-dideoxyinosine (DDI) in 40% of sequenced viruses in the absence of APOBEC3G or A3F but not against both at the same time [76].

During Vif APOBEC3G interaction, Vif K22 residue plays an important role in neutralizing APOBEC3G while its mutant form Vif K22H demonstrates declined effectiveness and elevates the amount of drug resisting mutation in the viral genome [89,94]. This refers to adaptation of HIV-1 virions in order to escape antiretroviral drugs, by tempering APOBEC3G due to high mutational load [95]. Currently researchers investigating methods to diminish the activity of Vif as a tempering APOBEC3G due to high mutational load [95]. Currently researchers investigating methods to diminish the activity of Vif as a tempering APOBEC3G due to high mutational load [95].

Effects of RT and APOBEC3G on Base Composition of HIV-1

According to a comparative study done on viral isolates (1998–2009) of HIV-1 Groups M, N and O did not show any major change in their nucleotide composition and remained conserved during evolution [98]. HIV-1 is a member of lentivirus family, which consists of A rich and C deficient RNA genome [99-101]. Another study also suggests that CpG content is actively suppressed in these viral genomes [102]. While comparing base of different HIV-1 viral genes such as gag, pol and env, constancy in the base composition still remains. It was also reported that comparatively pol gen is mostly A rich with gag slightly A deficient; while Rev, Tat and Nef contain a lower percentage of A nucleotide [99,100,103]. Such typical conservation in nucleotide composition might be related to RNA structure stability [104]. In a recent study HIV gag and pol genes were modified in such a manner that nearly 4% of A nucleotides, without affecting encoding of amino acids or disturbing sequences and AU rich elements important for Rev [104]. This resulted in decrease in cDNA synthesis and increase in stability of RNA. Further reduction in A content of pol showed comparatively more effect than gag gene on RNA genome. These results suggest a link between base composition and viral pathogenicity.

Conclusion

HIV-1 strictly is highly optimized to adapt for its survival, factors responsible for endangering its existence is removed by modification in its genome during its replication. RT and APOBEC3G play major function in modeling of HIV-1 genome and its diversity with the help of recombination and mutation. A better understanding of exact details related to the driving force behind the mutation and evolution of the HIV-1 genome is essential for the design of more efficient anti-retroviral therapies. Current studies suggest that key for survival of HIV-1 might be related to its nucleotide composition. Base composition with A nucleotide rich and C nucleotide deficient content is a distinct property of the members of lentivirus family. Richness of A and deficiency of C levels might be related to Structural flexibility and complexity to viral genome as they play a key role in drug resistance and adaptation. It is also important to notice here that APOBEC3G mediated G→A mutation mostly occurs in GC and GT. Here we would like to propose that a possible cure for HIV-1 could be achieved by increasing the amount of Thymine/Uracil (T/U) in nucleotide base composition. As already shown by Keating et al. in her research that decrease in A nucleotides lead to increase in stability of viral genome and also a decrease in its virulence and pathogenicity [105]. U/T nucleotides forms double bond with A during secondary structure formation. This will give rigidity to structure of Viral RNA. Firmness in structure of HIV-1 virus could also help ARV drugs to act more effectively as most drug resistance is the result of change in viral structure. Since G→A mutation does not target GT it will also help in decreasing complexity of viral genome and lead to decrease in drug resistance. There is still much to be discovered about distinct aspects related to Vif-APOBEC3G regulation, hotspot of HIV-1 mutations and different modes for overcoming viral multiple drug resistances for development of a proper anti-HIV-1 vaccine or drug therapy. It is equally necessary to specify various aspects related to reverse transcription, which checks its activity during mutation and recombination and could be utilized as a possible target for antiretroviral drugs. Extensive study of nucleotide content of HIV-1 could provide critical information related to completely stabilizing or un-stabilizing viral genome and lead to discovery of a better cure.

References


cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature 424: 94-98.


