Chapter 6 Genetic and Functional Characterization of HIV-1 VprC Variants from North India

Introduction

Five years after its discovery in North America, HIV was first reported in 1986 in India among sex workers in Chennai (Simoes et. al., 1987). The total estimate according to UNAIDS 2008 about people living with HIV in India was 0.36% that is 2.47 million people (UNAIDS, 2008).

The understanding of the mechanisms of natural protection from or spontaneous resistance to viral infections may contribute to the development of effective preventive measures. Also studies towards understanding the evolutionary history of the virus are important. Together such studies can reveal valuable insights for the design of effective vaccines. Several national and international agencies have worked towards the collection of data regarding the incidence and prevalence of HIV in the different states of our country.

NACO (National AIDS Control Organisation) a government body, NIHFW (National Institute of Health and Family Welfare), ICMR (Indian council of Medical Research) are few of the national organizations working in collaboration with several international bodies such as ILO (International Labour Organisation), WHO (World Health Organisation), UNAIDS to work out the future projections as well as control strategies. The combined effort of the above mentioned national organizations with international bodies like UNAIDS provides the direction in which various molecular virology studies can be carried out by various researchers. Research is essential to keep a track on the spread of the virus and on the variations in their sequences to design effective intervention strategies.
HIV diversity is associated with the error prone reverse transcriptase enzyme. The major subtypes of HIV identified are A-K depending on their geographical distribution. While subtype A is prevalent in Central Africa, subtype B is predominant in North and South American continent. The Indian subcontinent has subtype C as the major prevalent subtype along with China and South Africa, which is the major subtype contributing to the epidemic worldwide. Although, there is no direct correlation between the fitness of various subtypes to their respective spread, the prevalence of subtype C among majority of the HIV infected individual throughout the world would suggest that it has been naturally selected over other subtypes. The transmission rate of subtype C HIV-1 long terminal repeats (LTRs) from mother to infant has been shown to be a more likely event when compared to either subtype A or intersubtype recombinant LTR (Blackard et al., 2001). Also, in a recent review by scientists at NARI (National AIDS Research Institute), subtype C was stated to have fitness advantage over subtype A (Lakhashe et al., 2008).

In lieu of these findings, it is important to carryout separate studies in relation to subtype C. Major Studies pertaining to HIV-1 pathogenesis have been carried out on Subtype B from Europe and America. In contrast little information is available for subtype C, responsible for majority of the global epidemic (Africa and Asia). As far as the Indian subcontinent is concerned there are numerous reports available for the presence of subtype C throughout India (Sahni et al., 2002; Mandal et al., 2002, Kurle et al., 2004).

Less prevalent HIV-1 subtypes, other than subtype C have also been reported by several investigators in different parts of the country. These include subtype B and A from New Delhi and Punjab (Northern India), West Bengal and Manipur (eastern and north-eastern region), Maharashtra, Andhra Pradesh, Tamil Nadu and Gujarat (southern, western, and south western)
Characterization of HIV-1 VprC Variants | Chapter 6

(Gadkari et al., 1998; Kumar et al., 2006; Sidappa et al., 2004; Jameel et al., 1995, Cassol et al., 1996; Tripathy et al., 2007; Deshpande et al., 2009). Several recent reviews have been published by different groups describing the HIV epidemiology in India (Singh et al., 2008; Neogi et al., 2009).

The pure HIV subtypes are gradually phasing out and are being replaced by mosaic genome viruses owing to the recombinogenic nature of the virus (Levy et al., 2004; Shriner et al., 2004; Suryavanshi et al., 2007). Recent studies suggest that HIV proviruses can coexist in infected cells, which enables the formation of heterozygous virions and recombination may occur, inducing genomic diversification (Chen et al., 2005; Rhodes et al., 2003). Many recombinants between subtype A and C and between A and D have been reported from Eastern African countries which include Tanzania, Zambia, Uganda and Kenya, where subtype A, C and D usually circulate (McCutchan et al., 2000; Salminen et al., 1997).

The prevalence of unique recombinant forms (URFs) and circulatory recombinant forms (CRFs) were also reported from different part of India. URF B/C was reported from mainly north eastern part (Manipur) with different ancestral origin and breakpoint and B/C/D recombinants from Punjab (Lole et al., 1999), indicating that these URFs and CRFs may evolve through mixing of the co-existing strains or enter through trafficking due to the close geographical proximity. A report from north India showed the prevalence of multiple subtype including B’ (Thai B), C and A indicates that the genetic diversity of HIV in India is gradually expanding (Maitra et al., 1999). The A/C recombinant was identified from Maharashtra in India, where subtype C and A cocirculate (Deshpande et al., 2004).

The prevalence of circulating recombinant forms (CRFs) and unique recombinant form (URFs) indicates the changing patterns of HIV-1
epidemiology in India. The nature of HIV-1 circulating in different parts of India has been shown so far in the context of specific HIV-1 structural genes env and gag (Shankarappa et al., 2002; Sengupta et al., 2005a; Sengupta et al., 2005b).

Information pertaining to accessory genes from India also reveals the presence of multiple subtypes. There have been recent publications where other viral genes were studied for sequence variations among the existing subtypes in a process to identify hot spots for virus recombination. Banerjea and co-workers at N.I.I., New Delhi, recently analyzed sequences of HIV-1 LTR region from six HIV-1 infected individuals from North India to determine the nature and extent of variations (Neogi et al., 2009). HIV-1 LTR contains a number of cis-acting elements that are essential for proviral DNA synthesis, its integration into the host genome, transcription of viral genes and modulation of their expression (Millow et al., 2003). Neogi et al., found that four out of the six samples studied were close to subtype B forming a unique phylogenetic cluster. The remaining two samples (A3 and S3) turned out to be novel mosaic recombinants which showed resemblance with subtypes B, B/C-India and B/C – Myanmar gene segments.

HIV-1 Vpr, another viral accessory gene plays an important role in virus replication and can influence several viral and cellular functions (Engelbrecht et al., 2009). The purpose of carrying out this study was to find out the levels of genetic changes that are present in the Vpr gene from HIV-1 infected individuals from North-India (Punjab) and determine the functional relevance with respect to the two known important functions of this protein, its ability to activate HIV-1 LTR promoter and cause apoptosis. We recently reported subtype-B and C specific differences in Vpr mediated functions (Bano et al., 2007). Although genetic variations among several HIV-1 genes like Nef, Env and LTR have been reported from different regions of India.
(Jere et al., 2004; Rodriguez et al., 2007; Bhanja et al., 2007), this is the first comprehensive genetic/functional analysis of Vpr genes from Punjab region of North India which showed involvement of other subtypes in shaping the epidemic in this region.

**Experimental Procedure**

**Patient Information**

We obtained peripheral blood mononuclear cells from eight HIV-1 infected individuals (S1 - S3 and A4 - A6) from North India. These patients were registered for our study after obtaining requisite clearances. The patient information along with clinical parameters is summarized in Table 6-1. They were monitored at the Immunodeficiency Clinic of Post Graduate Institute of Medical Research and Education, Chandigarh, by our Clinician A. Wanchu. Peripheral blood mononuclear cells (PBMCs) were collected as described before (Husain et al., 1998).

<table>
<thead>
<tr>
<th>Subject Code</th>
<th>Age</th>
<th>Sex</th>
<th>Mode of Transmission</th>
<th>CD4 Count</th>
<th>CD8 Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>NII-PGI-IND-S1</td>
<td>33</td>
<td>M</td>
<td>Heterosexual</td>
<td>364</td>
<td>1543</td>
</tr>
<tr>
<td>NII-PGI-IND-S2</td>
<td>37</td>
<td>M</td>
<td>Heterosexual</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>NII-PGI-IND-S3</td>
<td>35</td>
<td>F</td>
<td>Heterosexual</td>
<td>253</td>
<td>1554</td>
</tr>
<tr>
<td>NII-PGI-IND-A4</td>
<td>30</td>
<td>M</td>
<td>Heterosexual</td>
<td>351</td>
<td>1872</td>
</tr>
<tr>
<td>NII-PGI-IND-A5</td>
<td>35</td>
<td>F</td>
<td>Heterosexual</td>
<td>387</td>
<td>2000</td>
</tr>
<tr>
<td>NII-PGI-IND-A6</td>
<td>28</td>
<td>F</td>
<td>Heterosexual</td>
<td>1046</td>
<td>1064</td>
</tr>
<tr>
<td>NII-PGI-IND-A7</td>
<td>35</td>
<td>M</td>
<td>Heterosexual</td>
<td>519</td>
<td>2000</td>
</tr>
<tr>
<td>NII-PGI-IND-A8</td>
<td>39</td>
<td>M</td>
<td>Heterosexual</td>
<td>447</td>
<td>1346</td>
</tr>
</tbody>
</table>

ND: Not Done.

**Genomic DNA isolation and amplification of Vpr gene**

PBMCs isolated from the patients were used to extract genomic DNA using Qiagen genomic DNA extraction kit. Isolated DNA was used to amplify sequence spanning the vpr gene by polymerase chain reaction. All PCRs were
performed with high fidelity Taq polymerase. The following primers P1 and P2 were used:-

P1-5' gc **GGATCC** ATGGAAACAAGCCCGAGAA 3' and

P2-5' gc **CTCGAG** CTTGGATCTACTGGCTCATTTC 3'.

*BamH1* and *XhoI* restriction sites (underlined) were engineered at the end of forward and reverse primers respectively for cloning into pcDNA3 expression vector (Invitrogen) under the CMV promoter for cellular studies. The positive clones were identified by the appearance of 291 bp insert after digesting with *BamH1* and *XhoI* restriction enzymes (Figure 6-1) and then further confirmed by sequencing. Using the same primers, we had earlier amplified the full-length Vpr genes from pNL4-3 DNA (Adachi *et al.*, 1986) and 93IN905 genetic subtype C clone (Lole *et al.*, 1999) as described in Chapter 4 (Figure 4-2). They were used as controls for all our cellular studies. Also point mutations were inserted in the VprC backbone to generate point mutants for performing functional transactivational studies. Leucine at the 22nd position was changed to Alanine and Leucine at the 64th position to Proline to identify the role of these variations observed in the Vpr protein isolated from the patient samples.

![Figure 6-1: Cloning of Vpr gene from patient DNA](image-url)
Sequencing, Computer alignment and analysis of vpr gene

Forward and reverse sequencing was performed with T7 and SP6 universal primers. Sequence alignment was carried out by ClustalW software (http://www.ebi.ac.uk/Tools/clustalw2/index.html) and edited manually to compare with pNL4-3 (AF324493) and 93IN905 (AF067158) prototype sequences. The bootscan analysis was performed in SimPlot ver 3.5.1 using Kimura (2-parameter) (Salminen et al., 1995). The ancestral and reference sequences were retrieved from HIV database located in Los Almos (http://www.hiv.lanl.gov).

Cell line

HEK 293 T cells were maintained in Dulbecco’s modified Eagle’s medium (Invitrogen) with 10% fetal bovine serum (Biolological Industries).

Western blot analysis

After 48 hours of transfection of 293 cells with indicated amounts of VprB, VprC and the three Vpr gene variants from HIV-1 infected individuals (VprS1, VprS2 and VprS3), cell lysates were prepared and protein content was estimated. 12% SDS-PAGE was used to resolve the proteins and they were transferred to the nitrocellulose membrane (BIORAD) at 300 mA for 1 hour. Polyclonal Vpr rabbit antiserum (Cat# 3951, NIH, MD, USA) was used as the primary antibody and HRP conjugated Rabbit IgG (Santa Cruz, CA, USA) was used as the secondary antibody. The blot was developed using ECL reagent (Amersham).

HIV-1 LTR promoter activation with Vpr gene variants

0.5 million HEK-293 cells were transfected with 100ng of HIV-1 LTR luciferase reporter construct along with the indicated amounts of wild type VprB-pcDNA3, VprC-pcDNA3 and the three Vpr gene variants cloned in pcDNA3. Twenty four hours post transfection, cell lysate was prepared using Reporter
Characterization of HIV-1 VprC Variants | Chapter 6

Lysis buffer (Promega) and luciferase activity was measured by luminometer as described in materials and methods.

**Apoptosis**

293 cells were transfected for 48 hours with the plasmids encoding wild type VprB, VprC and the three Vpr genes derived from HIV-1 infected individuals. Thereafter, cells were stained with Propidium-iodide (10 µg/ml) and flow cytometry was performed to determine apoptosis. FACS data were analyzed using WINMDI program.

**Results**

Figure 6-2: Comparison of Vpr gene sequences from the samples with consensus subtype B and C sequences

**Sequence analysis**

Sequences from all the 8 samples were aligned with subtype B (pNL4-3) and Indian subtype C (93IN905) sequences and the results are shown in Figure 6-2. The location of the three predicted α-helix is indicated at the top of the sequences. One of the most conserved and remarkable feature was the substitution of Leucine-64 with Proline (L64P) in 5 out of 8 samples,
suggesting strong evolutionary pressure for this change. Besides, sample S1 showed R32S and sample A5 showed W38R changes. The rest two samples, S2 and A6 were unique recombinants and a detailed genetic characterization is depicted in Figure 6-2.

**Phylogenetic analysis of Vpr variants**

Phylogenetic analysis of all the NII-PGI-IND-Vpr samples showed that 6 out of the 8 samples (S1, S3, A4, A7, A8 and A5) that possessed L64P mutation formed a separate cluster which segregated with the consensus and ancestral subtype C sequences. VprS2 formed a separate cluster and was positioned between C and B but sample A6 clearly showed branching with consensus B and D cluster (Figure 6-3) (shown by an arrow).

![Figure 6-3: Phylogenetic analyses](image-url)
Genetic characterization of the recombinants

In order to characterize the nature of recombinants the latter two samples (S2 and A6), were subjected to bootscan and phylogenetic analysis using SimPlot ver 3.5.1 and the result for sample S2 is shown in Figure 6-4.

Three break points located at 55, 86 and 135 nucleotide position, were identified. The N-terminal half consisted of B/D (phylogenetically close to B
France and ancestral (Anc) D, C (close to C India) and D (close to Anc D), while the C-terminal half consisted of C (close to Anc C). The sample A6 was also a mosaic of B/D, C, D and B subtypes. Three unique break points at nucleotide positions 54, 104, and 132 were identified. The N-terminal half consisted of B/D, C and D as observed with sample S2 but the C-terminal half was derived from B subtype (phylogenetically close to B France) (Figure 6-5).

Figure 6-5: Bootscan analysis for Sample A6
We carried out similar type of genetic analysis with one of the L64P samples (VprS3) and the only VprC sample that did not possess L64P mutation (VprA5). This showed almost 100% clustering with ancestral C (Figure 6-6) when subjected to bootscan analysis.

![Figure 6-6](image)

**Figure 6-6: Bootscan analyses for Samples S1 and A5**

**Rates of accumulation of non-synonymous and synonymous substitution**

According to the neutral theory of evolution, the number of synonymous substitutions with synonymous site (dS) is proportional to the rate of nucleotide mutation of a gene. The non-synonymous/synonymous ratio of rate constants (dN/dS) is indicative of the selection pressure at the protein level: dN/dS < 1 is indicative of purifying selection and amino acid conservation because of structural and functional constraints, and dN/dS > 1 is indicative of diversifying, positive selection where amino acid substitutions confer an advantage (Ganeshan et al., 1997). Six non-recombinants and two recombinants (VprS2 and A6) were subject to dN/dS analysis against Indian
Characterization of HIV-1 VprC Variants | Chapter 6

Seroconverter 93IN905 (Table 6-2). All eight samples showed purifying selection owing to structural and functional constraints.

Table 6-2: \( \frac{dN}{dS} \) ratio of Vpr variants

<table>
<thead>
<tr>
<th>Isolate</th>
<th>( \frac{dN}{dS} ) Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vpr S1</td>
<td>0.27</td>
</tr>
<tr>
<td>Vpr S2</td>
<td>0.19</td>
</tr>
<tr>
<td>Vpr S3</td>
<td>0.27</td>
</tr>
<tr>
<td>Vpr A4</td>
<td>0.28</td>
</tr>
<tr>
<td>Vpr A5</td>
<td>0.55</td>
</tr>
<tr>
<td>Vpr A6</td>
<td>0.15</td>
</tr>
<tr>
<td>Vpr A7</td>
<td>0.29</td>
</tr>
<tr>
<td>Vpr A8</td>
<td>0.29</td>
</tr>
</tbody>
</table>

HIV-1 LTR promoter activation is significantly less in Vpr samples with L64P mutation but not with recombinants

To find out how these changes in the Vpr gene impacted on the two well-known functions of the Vpr protein, namely, HIV-1 LTR promoter activation and apoptosis as described by us earlier (Bano et al., 2007), Luciferase assay and Propidium iodide staining were performed. Three samples were selected, two non-recombinant subtype C Vpr genes (S1 and S3), both containing L64P mutation, and one B/C/D recombinant (VprS1) for this purpose. They were cloned into pCDNA3 (Promega Biotech) under the influence of a powerful CMV promoter. Vpr encoding DNA was co-transfected along with indicated amounts of HIV-1 LTR-B Luciferase plasmid DNA as described in materials and methods and luciferase activity was measured. Representative results from 3 different experiments are summarized in Figure 6-7. Samples with proline in the 3rd a-helix (S1 and S3), were severely compromised with respect to their ability to transactivate HIV-1 LTR promoter. On the other hand
sample S3, which was a BD/C recombinant, retained the ability to transactivate HIV-1 LTR promoter activity. Control VprB and C showed about 2.5 fold more activation than LTR-B reporter gene alone. Also point mutant created in the backbone of subtype C Vpr gene derived from 93IN905 (L64P and L22A) were checked for their ability to transactivate LTR-B. While L64P showed a drastically reduced potential to transactivate LTR, the L22A mutant, on the contrary, was able to show dramatically enhanced transactivation potential.

![Graph](image-url)

Figure 6-7: Transactivation properties of various Vpr variants determined by Luciferase assay

**HIV-1 Vpr variants possess ability to cause apoptosis**

We then wanted to know what effect these three Vpr variants (VprS1, S2 and S3) would have on their ability to cause apoptosis in 293 cells as described earlier (Ranjan et al., 2007). Representative results from 3 independent experiments are shown in Figure 6-8. Percentage apoptosis with indicated Vpr variants is represented in the top right panel. As reported earlier, Vpr derived from an Indian C isolate (93IN905) (panel C) showed equivalent amounts of apoptosis as seen with VprB (panel B). All the three samples (VprS1- panel D, S2- panel E and S3 – panel F) were capable of inducing...
Characterization of HIV-1 VprC Variants | Chapter 6

apoptosis to varying extent that was significantly more than what was observed with untransfected cells (panel A). The numerical value at the right hand corner of each of the panels of Figure 6-8 denotes the percentage of apoptotic cell population.

![Figure 6-8: Apoptotic abilities of Vpr gene isolated from various samples](image)

**HIV-1 Vpr variants display similar intracellular stability**

It may be argued that Vpr variants (VprS1-3) with L64P mutation fail to transactivate HIV-1 LTR promoter because of their short half life of the protein itself or inefficient intracellular synthesis compared to control VprB or VprC proteins. To check this, the three variants along with VprB and VprC as control were transfected into HEK293 cells. Forty-eight hours post transfection the cells were lysed and divided into two equal proportions. One half was used to detect Vpr protein and the other for control actin protein. Equal amounts of protein from total cell lysates were subjected to gel analysis. There was no significant difference in the levels of Vpr protein when analyzed by western blot analysis using antibody specific for Vpr protein and the same
was true for the control actin protein (Figure 6-9). Control cells did not show any Vpr specific protein band.

Figure 6-9: Expression of Vpr protein from variants

Recombination in HIV-1 can be presumably thought to be the driving evolutionary force which helps the virus to adapt itself to varying selection pressures such as antiretroviral therapies. For recombination to occur, presence of multiple strains or multiple infections is essential. Earlier studies have negated the concept of multiple infections. But more recently the existence of multiple infections has been conclusively proven by Templeton and co-workers in 2009. The existence of multiple subtypes of HIV in a single
individual would lead to mixing of different lineages, ultimately leading to the evolution of URFs and CRFs. Recombinant viruses have already contributed substantially to the global pandemic, and the likelihood of generating recombinant viruses will continue to increase as the different HIV-1 subtypes spread worldwide (Peeters et al., 2000). There have been numerous reports for the presence of different subtypes in India in which different viral genes of different subclades exist together to produce a recombinant virus but the presence of sequences belonging to different subtypes within the same viral gene has been reported recently by Banerjea and co-workers at N.I.I., New Delhi.

In the present study 291 bp long Vpr gene was subjected to extensive genetic characterization studies. Several interesting features were highlighted, such as the presence of various subtypes namely B, C and D was observed in this region. Thus it could be deduced that Vpr gene sequence may play a crucial role in deciding the evolutionary fate of the virus. It might serve as a potential hot-spot for recombination.

Among the samples analysed, majority of them (6 out of 8) showed closeness to subtype C, but two of them were novel recombinants (S2 and A6). Most of them had a conserved mutation of L64P suggesting it to play a major role in HIV pathogenesis. The L64P mutants (VprS1 and VprS3) had a significantly reduced transactivation potential when compared to transactivation achieved with the control LTR-B reporter construct. This change is present almost in the middle of third α helix which is very likely to interfere with the helix formation and the functions associated with this region. It is possible that L64P mutants may act as a dominant negative factor with respect to HIV-1 LTR transactivation. However these L64P mutants showed normal competency in causing apoptosis. This is not surprising as others have reported similar observations with selective VprB mutants (Thotala et al.,...
Another conserved feature in all the samples except HIV infected individual A6 is the presence of 77Q which was earlier shown to be associated with Long term non-progression (Lumm et al., 2003). However, most of the patients showed normal progression to HIV/AIDS in our case. The two novel recombinants (S2 and A6) showed increased transactivation potential. Also, these recombinants possessed point mutations in the leucine repeat present in the 1st α helix of Vpr gene. This repeat region is essential for the transactivation abilities of Vpr.

The decreased potential to activate HIV-LTR promoter was not due to different amounts of Vpr variants being synthesized in the mammalian cells as determined by western blot analysis. None of the three samples (VprS1, S2 and S3) showed impairment of apoptotic potential of Vpr variants when compared with prototype Vpr B or C proteins. This is not surprising because the peptide containing the motif (Vpr 71-96) was fully conserved in all our Vpr variants. This peptide is known to induce apoptosis via a direct effect on the mitochondrial permeability transition pore (Jactot et al., 2000). Other mechanisms of Vpr mediated apoptosis have also been proposed (Andersen et al., 2006) and may be modulated to varying extent by our variants but that was not explored in this study.

We also suggest that these recombinations occurring within the gene are not random but predictable as the breakpoints relate to the points which have low entropy values (Figure 6-10). The virus has to selectively accumulate the functionally active recombinations which do not hamper the normal pathogenesis and this hypothesis has been supported by experiments carried out using HIV-1 env gene by different groups (Fan et al., 2007; Simon-Loriere et al., 2009). Thus, there exists a functional constraint over recombination events occurring within these small gene segments.
Such extensive mosaicism in a short stretch of 291nt encoding Vpr strongly suggests that this region not only is a potential hot spot for the formation of inter-subtype recombinants but also highlights the rapidly evolving HIV-1 epidemic in this region of India. This follows the global trend that show recombinants are being generated with great frequency and sometime they may be responsible for shaping the epidemic. These recombinants will impact on the choice of T-cell epitope based vaccines against HIV-1 for this region of India.