Global Variability of V3 Loop Tetrapeptide Motif: a Concern for HIV-1 Neutralizing Antibodies-based Vaccine Design and Antiretroviral Therapy

Luciano Rodrigo Lopes*†, Antonio Carlos da Silva Junior†, Paulo Bandiera-Paiva†, Jorge Casseb‡

1Bioinformatics and Bio-Data Science Division, Health Informatics Department, Universidade Federal de São Paulo–UNIFESP, São Paulo, SP, Brazil; 2Institute of Tropical Medicine of Sao Paulo - University of Sao Paulo, Laboratory of Medical Investigation LIM-56 / Faculty of Medicine –USP

ARTICLE INFO

Original Article

Keywords: HIV-1, gp120, V3 loop, GPGR motif, Vaccine

INTRODUCTION

Human immunodeficiency virus type 1 (HIV-1) presents an elevated evolution rate, resulting in high genetic variability. A particular intrahost environment pressures the variants, starting a selective competition between HIV-1 strains [1,2]. The dominance of a specific strain over others usually generates resistant variants [1]. Mutant strains can resist antiretroviral drugs or escape from neutralizing antibodies, increasing their replications and persisting as dominant populations [3–6]. While HIV-1 can resist antiretroviral drug combinations due to the accumulated mutations in target genes [7], a specific point mutation can confer escape from fusion inhibitor drug or neutralizing antibodies [8,9].

The primary neutralizing antibody target is HIV-1 gp120 (glycoprotein 120) subunit, an essential part of the envelope located on the viral surface [10]. HIV-1 gp120 is also a target of antiretroviral therapy, the fusion inhibitor drugs, due to its essential role in viral entry [11]. A central hypervariable region, V3 loop on gp120, plays a critical role in cell fusion, binding on the host CCR5 coreceptor through a relatively conserved tetrapeptide motif (GPGR, comprehending amino acids 312-315 of gp120) in the crown of the V3 loop [11–13]. The relative degree of conservation of GPGR is explained by the need to preserve the functional role of the V3 loop that appears to be essential for virus infectivity [14]. On the other hand, although the GPGR motif has been considered a conserved region of the V3 loop, this motif presents significant variability [3,15]. Diverses studies associated the specific variants of GPGR with fusion inhibitor drug resistance, escape from neutralizing antibodies, or even with pathogenicity [16–20]. Altogether, the diversity of the motifs at the crown of the V3 loop spread worldwide has been proven to be an epidemiologic problem.

Moreover, the characterization of variants of GPGR is an essential step in designing vaccines based on anti-V3 neutralizing antibodies. Thus, the present study aimed to perform a comprehensive analysis of the frequencies of
HIV-1 gp120 V3 loop tetrapeptide motifs, mapping their distribution rates in the countries and continents. We further described the clinical importance of the most prevalent GPGR variants, such as antiretroviral drug resistance, antibody escape, and pathogenicity.

**MATERIALS AND METHODS**

Based on HIV-1 gp120 amino acid sequences, we inferred the distribution of the gp120 V3 loop crown motifs. Amino acids sequences were obtained from the HIV Sequence Database of the Los Alamos National Laboratory (LANL) [21]. Sequences from each HIV-1 subtype and recombinants from all continents were included in the analyses, encompassing 259,288 gp120 amino acid sequences (representing all available sequences obtained in August 2020), using HXB2 as the reference sequence (https://www.ncbi.nlm.nih.gov/protein/1906385). The protein sequences from HIV-1 gp120 are available from the HIV Sequence Database (https://www.hiv.lanl.gov/). The sequences can be obtained by country, continents, or entirely. The amino acid sequences were sorted by geographic regions and aligned using MUSCLE [22] through MEGA X software [23]. The motif sequences were extracted from position 312-315 of gp120 based-alignment. The motifs frequencies and the graphs were performed using Microsoft Excel. Geographical maps based on motif frequencies were constructed using API geochart. The pie charts encompass the five most frequent HIV-1 V3 motifs in each continent.

Additionally, we presented a comprehensive bar graph including the world ten most frequent V3 variant motifs per country, in which prevalence rates were superior to 0.5%. Those motif frequencies with prevalence inferior to 0.5% were gathered in a group denominated "others." A file containing detailed data with all V3 loop motif frequencies for each country and their respective subtypes is available, Supplementary Table 1, that can be requested via jommid@gmail.com. The clinical meaning of the prominent relevant motifs was based on previous studies, properly cited.

This research included secondary biological data with no possibility of individual identification. The Unifesp Research Ethics Committee was consulted and determined that this study did not need ethical approval.

**RESULTS AND DISCUSSION**

The HIV-1 global frequencies analysis (Figure 1A) highlighted the predominance of two prominent V3 motifs, GPGR and GPGQ, in the continents (Figures 1B and C). HIV-1 strains harboring GPGR motifs were frequent in America, Europe, and Oceania. The GPGQ motif was highly presented among HIV-1 strains from Africa, Asia, Middle East, and USSR former. Previous studies showed the GPGR motif and the HIV-1 B subtype association, whereas the GPGQ motif was associated with non-B subtype strains [15,24].

Our results showed 88.7% of HIV-1 strains harboring GPGR motif classified as B subtype, while only 2.5% of HIV-1 strains harboring GPGQ were B subtype (Figure 2). The first sequenced HIV-1 genome carried a GPGR motif, corresponding to the B subtype [25]. The oldest known HIV-1 strain, identified in Africa in 1959 [26], was associated with the D subtype and carried the GPGQ motif at the crown of the V3 loop [26,27]. These facts can suggest a long-term predominance of the GPGQ motif over GPGR in the African continent.

It makes sense to consider a dominant strain to perform a vaccine during an epidemic. Considering that the motif in the crown of the V3 loop was shown to induce a humoral immune response [28], a vaccine focused on neutralizing antibodies must include GPGQ-containing epitopes rather than GPGR. HIV-1 strains that carry the GPGQ motif appear to elicit a broader neutralizing antibody-based immune response than HIV-1 GPGR strains [29, 30]. While neutralizing anti-V3 antibodies against the GPGR motif has reduced capacity to bind HIV-1 strains carrying the GPGQ strain [30,31], anti-V3 antibodies-based immune responses elicited by strains with GPGQ motif favor the neutralization for both GPGR and GPGQ strains [29,32].

Despite the predominance of HIV-1 strains carrying GPGR or GPGQ motifs, each continent had some particularities. Considering the five most frequent motifs in each continent, we showed twelve variants (Figure 1A). Some countries in the American subcontinent, Africa, Asia, and Europe, had a high diversity of V3 motifs, presenting more than ten variants (Figure 3). On the other hand, more significant variability was not observed in the Middle East and the USSR's former countries due to the smaller sample size from these regions. Thus, one of the limitations of our study refers to the obtained dataset. Although the LANL database is recognized as one of the most important databases, harboring HIV-1 sequences from all over the world, we observed unbalanced data due to the capacity of some countries to provide a more significant amount of viral genome sequences while others provided smaller data sets.

In total, more than five hundred V3 motif variants were found (Supplementary Table 1). The high diversity of cocirculating strains can compromise the potential effectiveness of the vaccine. The complexity of V3 motifs distribution may occur due to the selective pressure exerted by the immune system and/or fusion inhibitor drugs, promoting the emergence of dominant strains.

While HIV-1 GPGQ strain can escape from neutralizing antibodies and resists fusion inhibitor drugs, such as vicriviroc [33–35], the GPGR motif favors anti-V3 antibody binding and fusion inhibition by antiretroviral drugs (Table 1). In addition to GPGR and GPGQ motifs, strains that harbor GPGK, the third most frequent motif found in the continents, are susceptible to
**HIV-1 V3 variant motifs hamper vaccine design and therapy**

HIV-1 V3 variant motifs hamper vaccine design and therapy capacity of the strain [36]. On the other hand, the GPGA motif reduces antibody neutralization [40] but confers a high fitness cost, altering the CXCR4 coreceptor binding capacity [41,42]. Mutations within the V3 motifs may impact the molecular processes associated with membrane fusion, such as host receptors binding, syncytium formation, or cell-to-cell transmission [43,44].

**Fig. 1.** Distribution of GPGR motifs among HIV-1 circulating strains in the continents (A). Pie charts representing the most frequent variant motifs resulted in 12 variants. The global motif frequencies analysis highlights the predominance of GPGR and GPGQ in the continents. Heatmaps show the distribution of GPGR (B) and GPGQ (C) in the world. The frequency rates and the pie chart graphs were performed using Microsoft Excel. Geographical maps were constructed using API geochart.

Besides the GPGR and GPGQ motifs, massively found in all continents, relevant motifs are also commonly observed in HIV-1 strains: GPGK, GPGS, APGR, GPPG, PGA, GLGR, GQGR, and GWGR. These sequences encompass the ten most frequent V3 motifs around the world, together with GPGR and GPGQ. The binding capacity of anti-V3 neutralizing antibodies can be reduced in the presence of GPGK, GAGR, GLGR, and GPGA [20,36,40,45–47], impairing the course of HIV-1 infection and affecting a vaccine design (Table 1). Contrarily, GWGR increases the neutralizing capacity of antibodies [48].

The GWGR motif is part of the complex set of V3 motif variants composed of South American HIV-1 strains (Figure 1A). HIV-1 GWGR strains are highly distributed in Brazil. Previous studies identified up to 46% of strains harboring GWGR motifs [49–51]. According to our results, the GWGR motif was included in 12.9% out of 5,302 HIV-1 Brazilian strains (Figure 3). Although some neighboring countries of Brazil also presented HIV-1 GWGR strains circulating among infected populations, those frequency rates are lower (Figure 3). GWGR motif provides decreased pathogenicity to HIV-1 in comparison with GPGR. The patients infected with HIV-1 GWGR strains showed lower AIDS-defining events than those infected with...
GPGR strains [48,52]. Furthermore, the avidity of anti-V3 antibodies becomes increased by the GWGR motif, contributing to slower disease progression [48,50]. Altogether, the GWGR motif confers high fitness costs to HIV-1 strains, reducing viral pathogenicity without a clear compensation for these costs. Similarly, APGR, TPGR, and GPGG motifs also decrease the pathogenicity of HIV-1, impairing the viral infective capacity, syncytium formation, and the capacity to bind CCR5 coreceptor [14,43,53,54]. However, further analysis must be performed to find which benefits the APGR, TPGR, and GPGG motifs could provide to HIV-1 strains or whether the presence of these defective motifs could contribute with some compensatory state to HIV-1 strains. The low-pathogenic strains, such as HIV-1 GWGR, appear to reduce the disruption of host organism structures or biological processes. Consequently, this lack of pathogenesis may favor the viral persistence and even its transmission. However, the high prevalence of HIV-1 GPGR/Q strains suggests a better adaptive state than those harboring other motifs.

Fig. 2. Pie charts representing the subtype of HIV-1 strains harboring GPGR and GPGQ.

Table 1. Clinical meaning of V3 loop crown motifs variants

<table>
<thead>
<tr>
<th>Tetramer</th>
<th>Mutation</th>
<th>Antibody binding/neutralizing</th>
<th>Reference number</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAGR</td>
<td>P313A</td>
<td>Reduced</td>
<td>45</td>
</tr>
<tr>
<td>GAGR</td>
<td>P313A</td>
<td>Reduced</td>
<td>46</td>
</tr>
<tr>
<td>GLGR</td>
<td>P313L</td>
<td>Reduced</td>
<td>47</td>
</tr>
<tr>
<td>GWGR</td>
<td>P313W</td>
<td>Increased</td>
<td>48</td>
</tr>
<tr>
<td>GPGA</td>
<td>R315A</td>
<td>Reduced</td>
<td>40</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Reduced</td>
<td>20</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Reduced</td>
<td>36</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Reduced</td>
<td>37</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Reduced</td>
<td>38</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Reduced</td>
<td>39</td>
</tr>
<tr>
<td>GPGQ</td>
<td>R315Q</td>
<td>Reduced</td>
<td>31</td>
</tr>
<tr>
<td>GPGQ</td>
<td>R315Q</td>
<td>Reduced</td>
<td>40</td>
</tr>
</tbody>
</table>

**Drug susceptibility**

<table>
<thead>
<tr>
<th>Tetramer</th>
<th>Mutation</th>
<th>Sensitivity</th>
<th>Reference number</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Susceptible to cenicriviroc</td>
<td>20</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Susceptible to maravirocin</td>
<td>36</td>
</tr>
<tr>
<td>GPGQ</td>
<td>R315Q</td>
<td>Resistance to vicriviroc</td>
<td>34</td>
</tr>
</tbody>
</table>

**Fitness cost/Pathogenicity loss**

<table>
<thead>
<tr>
<th>Tetramer</th>
<th>Mutation</th>
<th>Phenotypic trait</th>
<th>Reference number</th>
</tr>
</thead>
<tbody>
<tr>
<td>APGR</td>
<td>G312A</td>
<td>CCR5 binding defective</td>
<td>53</td>
</tr>
<tr>
<td>TPGR</td>
<td>G312T</td>
<td>Defective syncytium formation</td>
<td>14</td>
</tr>
<tr>
<td>GWGR</td>
<td>P313W</td>
<td>Reduced pathogenesis</td>
<td>50</td>
</tr>
<tr>
<td>GPGR</td>
<td>R315A</td>
<td>Defective to bind CXCR4</td>
<td>41</td>
</tr>
<tr>
<td>GPGA</td>
<td>R315A</td>
<td>Defective to bind CXCR4 and deficient in inducing apoptosis</td>
<td>42</td>
</tr>
<tr>
<td>GPGK</td>
<td>R315K</td>
<td>Reduced replication rates</td>
<td>36</td>
</tr>
<tr>
<td>GPGG</td>
<td>R315G</td>
<td>Defective syncytium formation</td>
<td>43</td>
</tr>
<tr>
<td>GPGG</td>
<td>R315C</td>
<td>CCR5 binding defective</td>
<td>54</td>
</tr>
</tbody>
</table>
Fig. 3. Distribution of V3 motifs among HIV-1 circulating strains in each country from all continents. The bars represent the most frequent variant motifs (frequency > 0.5% among circulating strains), resulted in 10 variants. Motifs < 0.5% were included “others” group. Abbreviations based on ISO 3166-1 alfa-2 code represent each country.
Lopes et al.

Whereas some motifs can resist fusion inhibitor drugs or escape from neutralizing antibodies, the high variability of V3 crown motifs promotes an epidemiological challenge impacting the antiretroviral therapy decision and HIV-1 vaccine development. Although studies involving GPGR motif variants have been performed since the beginning of the AIDS pandemic, we contributed by performing an extensive analysis that included a vast number of GPGR variants. Additionally, our study’s essential goal was to determine the distribution of the variants of the GPGR motif of the V3 loop in each country and continent. Therefore, considering the mRNA vaccine approach successfully applied against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a potential anti-HIV-1 mRNA-based vaccine can be performed, including variants according to the geographical distribution. Combining the distribution rates of the most prevalent variants of GPGR with their clinical implication can contribute to the vaccine design or the decision for fusion inhibitor drug-based therapy.

ACKNOWLEDGMENT

The authors thank Mr. Vitor Machado Tonini for manuscript revision and Universidade Federal de São Paulo-UNIFESP for all support. This research included secondary biological data with no possibility of individual identification. No funding was available for this study. The Unifesp Research Ethics Committee was consulted and determined that this study did not need ethical approval.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest associated with this manuscript.

REFERENCES


HIV-1 V3 variant motifs hamper vaccine design and therapy


53. Suphaphiphat P, Essex M, Lee T-H. Mutations in the V3 stem versus the V3 crown and C4 region have different effects on the binding and fusion steps of human immunodeficiency virus type 1 gp120 interaction with the CCR5 coreceptor. Virology. 2007; 360 (1): 182–90.
