Insights from the interfaces of HIV-1 envelope (ENV) trimer viral protein GP160 (GP120-GP41)

Christina Nilofer, Arumugam Mohanapriya*
School of Biosciences and Technology, Vellore Institute of Technology, Vellore Campus, Tiruvalam Road, Katpadi, Vellore, Tamil Nadu - 632014, India

Received on: 02 Dec 2020
Revised on: 31 Dec 2020
Accepted on: 02 Jan 2021

Keywords:
HIV-1,
Envelope,
Trimer,
Glycoprotein,
GP160,
GP120,
GP41,
Protein Interface,
Van Der Waals

ABSTRACT

The Human Immunodeficiency Virus (HIV-1) type 1 viral protein is a life threatening virus causing HIV/AIDS in infected humans. The HIV-1 envelope (ENV) trimer glycoprotein GP160 (GP120-GP41) is gaining attention in recent years as a potential vaccine candidate for HIV-1/AIDS. However, the sequence variation and charge polarity at the interacting sites across clades is a shortcoming faced in the development of an effective HIV-1 vaccine. We analyzed the interfaces in terms of its interface area, interface size, and interface energies (van der Waals, hydrogen bonds, and electrostatics). The interfaces were divided as dominant (≥60%) and subdominant (<60%) based on van der Waals contribution to total energies. 88% of GP120 and 74% of GP41 interfaces are highly pronounced with van der Waals energy having large interfaces with interface size (98±65 (GP120) and 73±65 (GP41)) and interface area (882±1166Å² (GP120) and 921±1288Å² (GP41)). Nevertheless, 12% of GP120 and 26% of GP41 interfaces have subdominant van der Waals energies having small interfaces with interface size (58±20 (GP120) and 27±9 (GP41)) and interface area (581±1605Å² (GP120) and 483±896Å² (GP41)). It was interesting to observe GP41 small interfaces with subdominant van der Waals are stabilized by electrostatics (r²=0.63) without hydrogen bonds (r²=0). However, GP120 small interfaces were found to have two fold more hydrogen bonds (r²=0.59) than electrostatics (r²=0.20). Therefore, our previous finding stating that small protein-protein interfaces rich in electrostatics holds true in case of GP41 whereas not with GP120 protein interfaces.

INTRODUCTION

Regardless of the remarkable efforts to develop a vaccine for HIV-1/AIDS has always been a great challenge over the last decade with disappointing results in clinical trials (Shin, 2016). The unsatisfactory clinical trial results from VaxGen’s AIDSVAXgp120 vaccine and MRKAd5 HIV-1 Gag/Pol/Nef have been discussed elsewhere (Adis Editorial, 2003; Überla, 2008). This could be due to the viral human molecular mimicry, protein structural architecture, viral protein mutation and glycosylation. Despite the serious biotechnological challenges there is always an amplified energy to synthesis ENV trimer spike protein. The reasonable efficacy shown in the Thai trail vaccine (RV144 - ENV-GP120, Gag and Pro) is promising (Rerks-Ngarm et al., 2009, 2013). Post Thai trail (RV144), the focus is on envelope (ENV) as a vaccine candidate. In addition, ENV GP160 with least homology is selected by performing a sequence
comparison between HIV and human proteome (Kangueane et al., 2008). GP160 ENV trimer spike glycoprotein has gained attention as a potential vaccine candidate in the recent years. Production of native like HIV-1 envelope trimer glycoprotein is a challenge in designing, developing and validating an effective vaccine from a biochemical, structural and immunological viewpoint (Sanders and Moore, 2017; Doores, 2015).

Ringe et al. investigated on the number of factors that are importantly influencing the design, stability and purification of native like HIV-1 envelope trimer glycoprotein. Alsalmi et al. used strept tag method to purify GP160 trimer protein and was resulted with cleaved, uncleaved, fully or partially glycosylated trimers. In addition, they found cleaved gp140 were not required for trimerization, however they played a significant role in triggering conformational changes in channelizing the trimers to generate compact three blade propeller shaped trimers. Verkerke et al. used lectin affinity chromatography to purify native like trimers from diverse HIV-1 isolates. The challenges faced in the production, analysis and synthesis of GP160 ENV trimer glycoprotein are reported (Grimm et al., 2015; Guenaga, 2015). Surface mutation, charge polarity and glycosylation and sequence variation between known variants in different clades are the significant barriers causing difficulty in imitating a native-like conformation of the glycoprotein. It is evident that assembling individual GP160 into a trimer spike complex structure is a challenge from a protein-protein interaction viewpoint. A large number of GP120 and GP41 structures are available in the PDB deposited using different biophysical techniques to understand the underlying molecular mechanism of the interacting proteins.

Sowmya et al. demonstrated the correlation between sequence polarity and mean Shannon entropy by calculating sequence polarity for surface residues in GP120 and GP41 and concluded stating the use of protein modification in the enhancement of HIV-1 vaccine across different clades, blood, and brain. Nilofer et al. characterized the interfaces of GP120-GP120, GP120-GP41 and GP41-GP41 and reported that the interfaces of GP120-GP120 are largely polar. The interfaces of GP120-GP41 and GP41-GP41 are characterized as polar and non-polar residues. We characterize a manually curated dataset of 121 GP120 and 85 GP41 (Figure 1) protein interfaces reported by Nilofer et al. using interface features including interface area, interface size (number of residues at the interface), van der Waals, hydrogen bonds and electrostatics, to verify our previous finding stating that small protein interfaces are rich in electrostatics are often linked to regulatory proteins (Nilofer et al., 2020). The residues at the interface are displayed using CPK depiction (Discovery Studio® (Systèmes, 2020).

**MATERIALS AND METHODS**

**Dataset**

We used a dataset of 206 interfaces manually curated as reported by Nilofer et al. It consists of 121 GP120 (Table 1) and 85 GP41 (Table 2) interfaces. It should be noted that GP120 structures in the PDB are available in ligand-bound state.

**Interface area**

Interface area was estimated for each of 121 interfaces of GP120 and 85 interfaces of GP41 using Naccess (Hubbard and Thornton, 1993). Naccess uses Lee and Richards method (Lee and Richards, 1971), wherein a probe with radius 1.4Å (Jones and Thornton, 1996) roll over the protein complex in monomer state and dimer state to find the accessible surface area and the interface area using delta ASA. Delta ASA (change in accessible surface area) is calculated using a formula: [ASA (Monomer subunit 1) + ASA (Monomer subunit 2) - AB (Dimer complex)]/2.

**Interface size & interface energies**

Interface size and interface energies were estimated for each of 121 interfaces of GP120 and GP41 using PPCheck (Sukhwal and Sowdhamini, 2015). PPCheck uses distance criteria to identify the non-covalent interactions between atoms of the two interacting proteins. It should be noted that the role of water is ignored in this analysis.

**Large interface area and small interface area**

Interfaces with large interface size (98±65 (GP120) and 73±65 (GP41)) and interface area (882±1166Å² (GP120) and 921±1288Å² (GP41)). Having dominant van der Waals energy (≥60%) at the interface are defined as large interface whereas interfaces with small interface size (58±20 (GP120) and 27±9 (GP41)) and interface area (581±1605Å² (GP120) and 483±896Å² (GP41)) having subdominant van der Waals energy (<60%) are defined as small interface.

**Dominant and subdominant van der Waals interface**

Interfaces with van der Waals contribution ≥60% to total energy (sum of van der Waals, hydrogen bonds and electrostatics) is defined as dominant interfaces, while interfaces with van der Waals contribution <60% to total energy is defined as subdominant interfaces. A cutoff of 60% was used as the larger part of van der Waals contribution was at...
Table 1: Dataset of GP120 (121) having large (van der Waals Dominant) and small (van der Waals Subdominant) interfaces are listed.

<table>
<thead>
<tr>
<th>Large interfaces van der waals dominant</th>
<th>GP120 Small interfaces van der waals subdominant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1G9M 2NY4 4R4N 1RZ7 3IDX 4J6R 4LSS 4R4H 4YBL 5IES</td>
<td>1G9N 2NY5 4RFN 1RZ8 3LQA 4JB9 4LSU 4RQS 4YCL</td>
</tr>
<tr>
<td>1GC1 2NY6 4RF0 1RZF 3NGB 4J03 4OLU 4RX4 4YDK</td>
<td>1NAK 3JWD 4XNZ 1RZG 3Q6G 4J03 4OLU 4RX4 4YDK</td>
</tr>
<tr>
<td>1RZJ 3JWO 4YDJ 1RZI 3SE8 4JK8 4OLV 4S1Q 4YDL</td>
<td>1RZK 3MLS 4ZTO 1YYL 3SE9 4JPV 4OLW 4S1R 4YFL</td>
</tr>
<tr>
<td>2F58 4DVR 5F96 1YYM 3TYG 4JPW 4OLX 4S1S 5CAY</td>
<td>2F58 4DVR 5F96 1YYM 3TYG 4JPW 4OLX 4S1S 5CAY</td>
</tr>
<tr>
<td>2NY0 4K0A 5TE7 2I60 4D9L 4LAJ 4OM0 4XMP 4F9W</td>
<td>2NY2 4LST 5TE7 2I60 4D9L 4LAJ 4OM0 4XMP 4F9W</td>
</tr>
<tr>
<td>2NY1 4KA2 5TE7 2I60 4D9L 4LAJ 4OM0 4XMP 4F9W</td>
<td>2NY3 4R4F 5TE7 2I60 4D9L 4LAJ 4OM0 4XMP 4F9W</td>
</tr>
</tbody>
</table>

Figure 1: Examples of GP120 (PDB 3MLZ) and GP41 (PDB 2ZFC) viral protein interfaces are illustrated along with the percent contribution of each of interface energies.

Statistical analysis

We calculated interface energies of GP120 and GP41 using the statistical (Microsoft® Office Excel (version 2003)) variables including mean, mode, distribution, standard deviation and frequency at definite bin and range. We also carried out multiple linear regressions analysis for each interface with interface size against van der Waals, hydrogen bonds, electrostatic, total energy and interface area using regression tool. Its co-efficient of determination ($r^2$) was predicted with an evaluation of p-value using ANOVA (statistical test) at 95% confidence limit.

RESULTS AND DISCUSSION

The HIV-1 envelope trimer glycoprotein GP160 is a potential vaccine candidate for HIV-1/AIDS (Burton et al., 2004). Structural data of GP160 available in
Table 2: Dataset of GP41 (85) having large (van der Waals Dominant) and small (van der Waals Subdominant) interfaces are listed.

<table>
<thead>
<tr>
<th>GP41</th>
<th>Large Interfaces van der Waals Dominant</th>
<th>Small Interfaces van der Waals Subdominant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1CE0</td>
<td>1U8N 2P8L 3DRT 3JWO</td>
<td>1NLD 3MA9 4NGH 5KA6</td>
</tr>
<tr>
<td>1TG</td>
<td>1U8O 2P8M 3EGS 3LEX</td>
<td>1TZG 3MAC 4NH 5MP6</td>
</tr>
<tr>
<td>1TH</td>
<td>1U8P 2P8P 3F4Z 3LEY</td>
<td>2CMR 3MNW 4UG 5U3K</td>
</tr>
<tr>
<td>1TJ</td>
<td>1U8Q 2PW1 3FN0 4NRX</td>
<td>2FX7 3040 4XAW 5U3L</td>
</tr>
<tr>
<td>1U8H</td>
<td>1U91 2PW2 3IDG 4XC1</td>
<td>2FX9 3043 4XBE 5U3M</td>
</tr>
<tr>
<td>1U8I</td>
<td>1U92 2ZFC 3IDI 4XC</td>
<td>2Q3I 3029 4XCF 5U3N</td>
</tr>
<tr>
<td>1U8J</td>
<td>1U93 3D0L 3IDJ 5FB9</td>
<td>2R5B 3P30 5CC 5U3O</td>
</tr>
<tr>
<td>1U8K</td>
<td>1U95 3D0V 3IDM 5U3J</td>
<td>2R5D 3UIA 5CMU -</td>
</tr>
<tr>
<td>1U8L</td>
<td>2F5B 3DRO 3IN -</td>
<td>2X7R 4KHT 5IQ7 -</td>
</tr>
<tr>
<td>1U8M</td>
<td>2FX8 3DRQ 3JWD -</td>
<td>3ECB 4KH 5IQ9 -</td>
</tr>
</tbody>
</table>

The PDB are always found to be coupled with ligand showing the degree of stability of GP160 without supportive ligand (Moore et al., 1992). The trimer interfaces are unstable when produced invitro and this may be due to its sequence composition and structural conformation (Moore et al., 1990; Chen et al., 2005). Nilofer et al. reported that the interfaces of GP120 are largely polar whereas the interfaces of GP120-GP41 and GP41 are characteristics of equal contribution of polar and non-polar residues. Interfaces with high polarity have an immense impact on protein's surface, immunological and stability properties. High polarity at the interface is bottleneck in the invitro synthesis and production of GP160 in a stable form. The instability of GP160 could also be due to the complicatedness in mimicking the invivo environment in invitro for protein folding and assembly of the complex (Abagyan and Batalov, 1997; Kinjo et al., 2001).

Figure 2: Mean interface size and interface area of GP120 and GP41 protein interfaces.

Therefore, we characterized the interfaces of GP120 and GP41 using interface area, interface size and interface energies using PPCheck (identifies non-covalent interactions using distance criteria). To verify our previous findings, we used the manually curated dataset of 121 and 85 interfaces of GP120 and GP41 proteins. The statistical analysis show that the mean interface size (98±65 (GP120) and 73±65 (GP41)) and interface area (882±1166Å² (GP120) and 921±1288Å² (GP41)) to be in close proximity for GP120 and GP41 interfaces (Figure 2). In contrast to our previous study we observed, most of the interfaces to have an interface area <1000Å² in both GP120 (60%) and GP41 (71%) and about 25% of GP120 and 19% of GP41 to have interface area between 1500Å² to 2000Å² (Figure 3).

Figure 3: Graph showing interfaces with interface area (Å²) for GP120 and GP41.

Subsequently, we described each interface of GP120 and GP41 using van der Waals, H-bond, electrostatics and total energies along with their varying proportion of contribution at the interface. Thus, we calculated each individual contribution in percentage towards total energy. We observed interfaces to have high percentage of van der Waals (77%) and a low percentage of hydrogen bonds (12%) and electrostatics (11%) on average for GP120 and GP41 complexes (Figure 4). In addition, we noticed the interfaces of GP120 and GP41 to be normally distributed with increasing percentage of van der Waals (Figure 5). While a proportion of the interface...
Figure 4: GP120 and GP41 interfaces are shown in terms of mean percent van der Waals, hydrogen bonds and electrostatics.

Figure 5: GP120 and GP41 interfaces are shown with varying percent contribution of van der Waals, hydrogen bonds and electrostatics.

decrease with increasing percentage of hydrogen bonds and electrostatics unlike van der Waals energy. It should be noted that interfaces of GP120 and GP41 are similar with the percentage contribution of van der Waals, hydrogen bonds, electrostatics, interface area and interface size. We further grouped the interfaces of GP120 and GP41 as dominant (≥60%) and subdominant (<60%) van der Waals based on its contribution towards total energy. As a result dominant interfaces have ≥60% of van der Waals with less magnitude of hydrogen bonds and electrostatics. We observed majority of interfaces of GP120 (88%) and GP41 (74%) to be van der Waals dominant with less that 10% contribution of hydrogen bonds and electrostatics while the remaining 12% of GP120 and 26% of GP41 have subdominant van der Waals with more than 15% of hydrogen bonds and electrostatic (Figure 6). We performed statistical analysis on dominant van der Waals and subdominant van der Waals interfaces to
Figure 6: GP120 and GP41 interfaces are shown with increasing percentages of van der Waals, hydrogen bonds and electrostatics.

Figure 7: Mean percentage of hydrogen bonds and electrostatics in large and small interfaces of GP120 and GP41 are shown.
highlight the contribution of hydrogen bonds and electrostatics in the small interfaces. We observed subdominant van der Waals interfaces of GP120 and GP41 to have three fold more of both hydrogen bonds and electrostatics when compared to dominant van der Waals interfaces (Figure 7). Furthermore, we noticed subdominant van der Waals interfaces to be more pronounced with more than 20% of hydrogen bonds and electrostatics distinct compared to dominant van der Waals interfaces (Figure 8). It is evident from (Figure 9) that the interface size and interface area of small interfaces are only half when compared to the large interfaces. Most of the interfaces with subdominant van der Waals have interface area less than 500Å² (Figure 10). Therefore, it was stated that the small interfaces with subdominant van der Waals energy and small interface area are rich in electrostatics. However, in context to the small interfaces (subdominant van der Waals) of GP120 and GP41, small interfaces of GP120 are rich in hydrogen bonds and GP41 is rich in electrostatics. Interface size increases with interface area is a known fact and hence we correlated interface size and interface energies. It was reported that total
Figure 10: Interface area in large and small interfaces among GP120 and GP41 is shown.

Figure 11: Correlation between interface energy and interface size of GP120 and GP41 interfaces is shown.
energy, van der Waals and hydrogen bonds increase with interface size but electrostats decrease with increasing interface size (Nilofer et al., 2020). While the results of our current study shows that van der Waals and total energies of GP120 and GP41 interfaces increase with interface size but hydrogen bonds and electrostats decrease with increasing interface size (Figure 11). Hence we divided our interfaces as large (dominant van der Waals) and small (subdominant van der Waals) interface based on their percent contribution towards total energy to check for correlation. It has also been reported that in small interfaces, total energy, van der Waals and hydrogen bonds decreases considerably with the increasing interface size whereas electrostats moderately increases with interface size (Nilofer et al., 2020). But, this is not the case with the small interfaces of GP120 and GP41. Surprisingly, we found electrostats (r²=0.63) (Figure 12p) to be highly pronounced in GP41 interfaces with subdominant van der Waals having van der Waals (r²=0.23) (Figure 12h) and without hydrogen bonds (r²=0) (Figure 12p) contribution. Contrastingly, we observed the small interfaces of GP120 to be highly stabilized by hydrogen bonds (r²=0.59) (Figure 12k) followed by electrostats (r²=0.20) (Figure 12o). Hence, we report that hydrogen bonds (r²=0.59) (Figure 12k) increases with the interface size in the small interfaces of GP120 and electrostats (r²=0.63) (Figure 12p) increases with the interface size in the small interfaces of GP41.

**CONCLUSIONS**

GP120 viral proteins interact with GP41 to form GP160 the HIV-1 trimer glycoprotein. Statistical analysis on the interfaces of GP120 and GP41 using interface area, interface size and interface energies including van der Waals, hydrogen bonds and electrostats demonstrate that they are similar. 88% of GP120 and 74% of GP41 interfaces have large interface area and interface size with dominant van der Waals energy; while 12% of GP120 and 26% of GP41 interfaces have small interface area and interface size with subdominant van der Waals energy. In addition, small interfaces were observed to have three fold more of hydrogen bonds and electrostats than large interfaces. It is shown hydrogen bonds to increase with interface size in the small interfaces of GP120; while electrostats to increase with interface size in small interfaces of GP41 in absence of hydrogen bonds. These insights from the interfaces of GP120 and GP41 shows that our previous finding stating that small interfaces with small interface area are rich in electrostats holds true in case of GP41 but not in the case of GP120.

**Funding Support**

The authors declare that they have no funding support for this study.

**Conflict of Interest**

The authors declare that they have no conflict of interest for this study.
REFERENCES


